



Relations entre le modèle d'image et le nombre de mesures pour une super-résolution fidèle

Yann Traonmilin

► To cite this version:

Yann Traonmilin. Relations entre le modèle d'image et le nombre de mesures pour une super-résolution fidèle. Traitement du signal et de l'image [eess.SP]. Télécom ParisTech, 2014. Français. NNT : 2014-ENST-0043 . tel-01135196

HAL Id: tel-01135196

<https://theses.hal.science/tel-01135196>

Submitted on 24 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Doctorat ParisTech
THÈSE

pour obtenir le grade de docteur délivré par

Télécom ParisTech
Spécialité : Signal et Images

Présentée et soutenue publiquement par

Yann Traonmilin

le 9 juillet 2014

**Relations entre le modèle d'image et le
nombre de mesures pour une
super-résolution fidèle**

Jury :

Jean-François Aujol, Professeur, Institut de Mathématiques de Bordeaux

François Malgouyres, Professeur, Institut de Mathématiques de Toulouse

Gabriel Peyré, Chargé de Recherche, CNRS CEREMADE, Université Paris Dauphine

Albert Cohen, Professeur, LJLL, Université Pierre et Marie Curie

Alexandre Gramfort, Maître de Conférence, CNRS LTCI, Télécom ParisTech

Mila Nikolova, Directrice de Recherche, CNRS CMLA, ENS Cachan

Andrés Almansa, Chargé de Recherche, CNRS LTCI, Télécom ParisTech

Saïd Ladjal, Maître de Conférence, CNRS LTCI, Télécom ParisTech

Rapporteur

Rapporteur

Rapporteur

Examineur

Examineur

Examinatrice

Directeur

Directeur

THÈSE

TELECOM ParisTech

École de l'Institut Mines-Télécom - Membre de ParisTech

46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

REMERCIEMENTS

A tous ceux qui m'ont aidé et soutenu pendant ces trois ans : Andrés et Saïd, ma famille, mes amis, mes collègues et surtout Scarlett.

Je remercie également les membres du Jury d'avoir accepté de lire mes travaux.

TABLE DES MATIÈRES

1	Introduction	7
1.1	Repliement spectral et limites matérielles	9
1.2	Problématique	11
1.3	Contributions	11
1.4	Publications	12
2	Méthodes variationnelles et super-résolution	13
2.1	Modèle de super-résolution	13
2.1.1	Acquisition d'une série d'images	13
2.1.2	Description des paramètres du modèle	15
2.2	Problèmes inverses d'imagerie	16
2.2.1	Problèmes convexes et non convexes	18
2.2.2	Le cas de la super-résolution	19
2.2.3	Modèles de régularité pour les images	20
2.3	Description formelle de la problématique	20
2.3.1	Super-résolution bien informée	20
2.3.2	Plan de la thèse revisité	21
3	Nombre d'images pour une super-résolution bien informée	23
3.1	Inversibilité de l'interpolation SR affine	23
3.1.1	Hypothèses	23
3.1.2	Théorème d'inversibilité	26
3.1.3	Résultats intermédiaires pour l'inversibilité	28
3.2	Conditionnement en fonction du nombre d'images	30
3.2.1	Interpolation	30
3.2.2	Estimation des paramètres de mouvement	38
3.2.3	Conclusion sur le conditionnement et expériences réelles	41
3.2.4	Calculs intermédiaires pour la hessienne	43
4	Application à la super-résolution à haute dynamique	45
4.1	Méthode de reconstruction	46
4.2	Stratégie pour le choix des paramètres d'acquisition	46
4.3	Borne d'erreur de reconstruction pour la HDR-SR	47
4.3.1	Nombre optimal d'images pour un temps d'exposition total fixe	47
4.3.2	Étude expérimentale de $f(N)$	48
4.3.3	Discussion et conséquences pratiques	48
4.4	Expériences réelles	48
4.5	Calculs intermédiaires pour la borne sur l'erreur de reconstruction	56
5	Régularisation des problèmes de SR inversibles mal conditionnés	57
5.1	Conditionnement local	57
5.2	Régularisation locale	62

5.2.1	Régularisation locale par troncature du gradient conjugué	62
5.2.2	Régularisation locale par pondération de la TV	63
5.3	Expériences	65
5.3.1	Application au dématricage multi-image	71
5.4	Conclusion	71
6	Robustesse de la super-résolution L^1 aux données aberrantes	73
6.1	Lien entre robustesse et reconstruction parcimonieuse	73
6.1.1	État de l'art	73
6.1.2	Définitions	74
6.1.3	Démonstration directe de la caractérisation par la PNC	76
6.1.4	Stabilité	77
6.2	Conditions de reconstruction parfaite par SR en présence d'outliers	78
6.2.1	Conditions suffisantes pour le K -pardon	78
6.2.2	Étude d'une structure d'outliers particulière	82
6.2.3	Algorithmes et exemples	83
6.3	Conclusion	84
7	Régularisation parcimonieuse et robustesse de la SR L^1	87
7.1	Extension du pardon au cas régularisé	87
7.1.1	Hypothèses de parcimonie sur les images naturelles	88
7.1.2	Définitions	88
7.1.3	Propriété de Reconstruction Exacte	90
7.1.4	Caractérisation par la propriété de non concentration	91
7.1.5	La Restricted Isometry Property pour le pardon régularisé	92
7.2	Le cas des matrices aléatoires	98
7.2.1	Revue des bornes existantes pour $\lambda = 1$	98
7.2.2	K, S -pardon en fonction de λ	99
7.2.3	Le cas d'observations aléatoires d'un signal parcimonieux	101
7.2.4	Conclusion sur les matrices aléatoires	102
7.3	Robustesse de la super-resolution L^1 -TV	102
7.3.1	Débruitage multi-image	102
7.3.2	Super-résolution	104
8	Conclusion	113
	Bibliographie	117

Abbreviations

- BR : Basse Résolution
- CG : Gradient conjugué (Conjugate Gradient)
- dB : Decibel
- HR : Haute Résolution
- PNC : Propriété de Non Concentration
- PRE : Propriété de reconstruction exacte (Exact Reconstruction Property)
- PSF : Flou de l'appareil photo (Point Spread Function)
- PSNR : Rapport signal sur bruit normalisé (Peak Signal to Noise Ratio)
- RIP : Propriété d'isométrie restreinte (Restricted Isometry Property)
- *t.q.* : Tel que
- SR : Super-Résolution
- TV : variation totale (Total Variation)

Définitions et notations

- \mathbb{N} : Ensemble des entiers naturels
- \mathbb{R} : Ensemble des nombres réels
- \mathbb{C} : Ensemble des nombres complexes
- $x \bmod y$: x modulo y
- I : Matrice identité
- $\dim V$: Dimension de l'espace vectoriel V
- $\text{Im} A$: Espace image de la matrice (où application linéaire) A
- $\ker A$: Noyau de la matrice (où application linéaire) A
- $\text{tr}(A)$: Trace de A
- $v.p.(A)$: Valeurs propres de A
- $\|u\|_0$: Pour $u \in \mathbb{C}^n$, $\|u\|_0 = \text{card}\{|u_i| \neq 0\}$.
- $\|u\|_q$: Norme L^q , pour $u \in \mathbb{C}^n \rightarrow \mathbb{C}$, $\|u\|_q = (\int |u|^q)^{1/q}$. Pour $u \in \mathbb{C}^n$, $\|x\|_q = (\sum_{i=1,n} |u_i|^q)^{1/q}$.
- A^H : Matrice transconjugée de A .
- A^T : Matrice transposée
- A^\dagger : Pseudo-inverse d'une matrice $A = (A^H A)^{-1} A^H$.
- $A_{T|}$: Si T est un support (vecteur de 0 et de 1). $A_{T|}$ est la restriction de la matrice A aux lignes indexées par T .
- B_T : B_T est la restriction de la matrice B aux colonnes indexées par T .
- $\mathcal{L}.\mathbb{R}^N$: L'ensemble des vecteurs de \mathbb{R}^N qui s'écrivent $x.L$, $x \in \mathbb{R}^N$, $L \in \mathcal{L}$. Le point est le produit coordonnée à coordonnée.
- **Transformée de Fourier** Soit u une image. Sa transformée de Fourier \hat{u} sur une grille Γ est

$$\hat{u}(\omega) = \sum_{\mathbf{x} \in \Gamma} u(\mathbf{x}) e^{-j\langle \omega, \mathbf{x} \rangle} \quad (1)$$

CHAPITRE 1

INTRODUCTION

La photographie d'une scène naturelle est un processus court, que l'on peut rarement reproduire. En effet, les conditions de lumière, la position des objets de la scène ou bien la position du photographe sont souvent difficiles à réunir de nouveau. Par exemple, on ne pourra reproduire une photographie de coucher de soleil prise pendant des vacances. Il faut donc être sûr de capturer tout les détails de la scène que l'on souhaite, car revenir prendre une autre photographie avec les mêmes conditions est impossible. D'ailleurs, une technique utilisée par les photographes professionnels est de toujours prendre une multitude de photographies de la même scène pour ensuite sélectionner la meilleure et ainsi s'assurer d'avoir une bonne image finale.

Ce côté éphémère des scènes à photographier est une raison qui pousse les fabricants d'appareils photographiques à augmenter la quantité d'information capturée par chaque photographie. Cette augmentation se traduit par l'amélioration de la résolution des photographies, c'est-à-dire du nombre de pixels par image enregistrée. Le photographe peut ainsi retoucher ses images après coup. Par exemple, une grande résolution permet de réaliser un gros plan de bonne qualité sur une partie de l'image. Une autre utilité d'une bonne résolution est la possibilité d'imprimer l'image en grand format pour des affiches. On présente figure 1.1, la différence entre un *zoom* (agrandissement) à partir d'une image de 0.5 Mégapixel et une de 0.125 Mégapixel. La qualité visuelle du zoom est meilleure avec la version de plus haute résolution. On observe un effet de crénelage sur le zoom à partir de l'image basse résolution. Cet effet s'appelle le repliement spectral (« aliasing ») et se produit lorsque la résolution n'est pas assez fine pour capturer correctement tous les détails de l'image.

Un problème survient lorsque l'on ne peut pas augmenter la résolution des capteurs, que ce soit pour des raisons économiques ou à cause de limites plus fondamentales pour les appareils haut de gamme. Avec de telles contraintes, comment augmenter la résolution des photographies ?

Une première façon est de prendre plusieurs photographies de la même scène en décalant l'endroit de la prise de vue pour échantillonner une information différente. On remarque ainsi que, si l'on connaît les mouvements relatifs, on augmente le niveau de détail capturé. Ce processus de capture de l'information est illustré sur la figure 1.2. On représente l'effet de prendre la même scène en utilisant plusieurs photographies de basse résolution (BR) avec des décalages pour chaque photographie. Les détails de la scène deviennent apparents. Produire une image de haute résolution (HR) à partir de ces images de basse résolution (BR) s'appelle l'imagerie de super-résolution. Prendre plusieurs photographies revient à échantillonner l'image sur une grille irrégulière. La super-résolution consiste donc à déterminer la position de chaque image puis interpoler l'image sur une grille régulière de haute résolution.

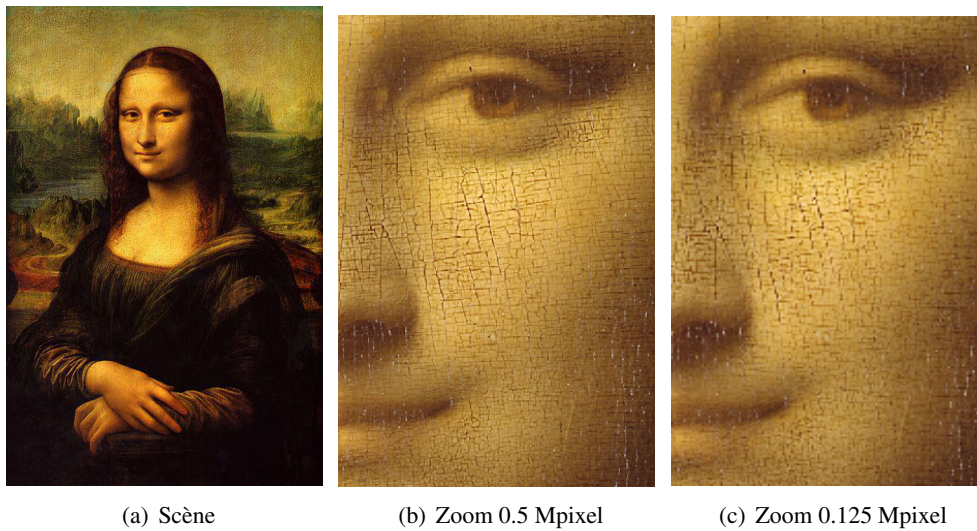


FIGURE 1.1. Utilité d'une bonne résolution en vue d'un traitement postérieur à l'acquisition.

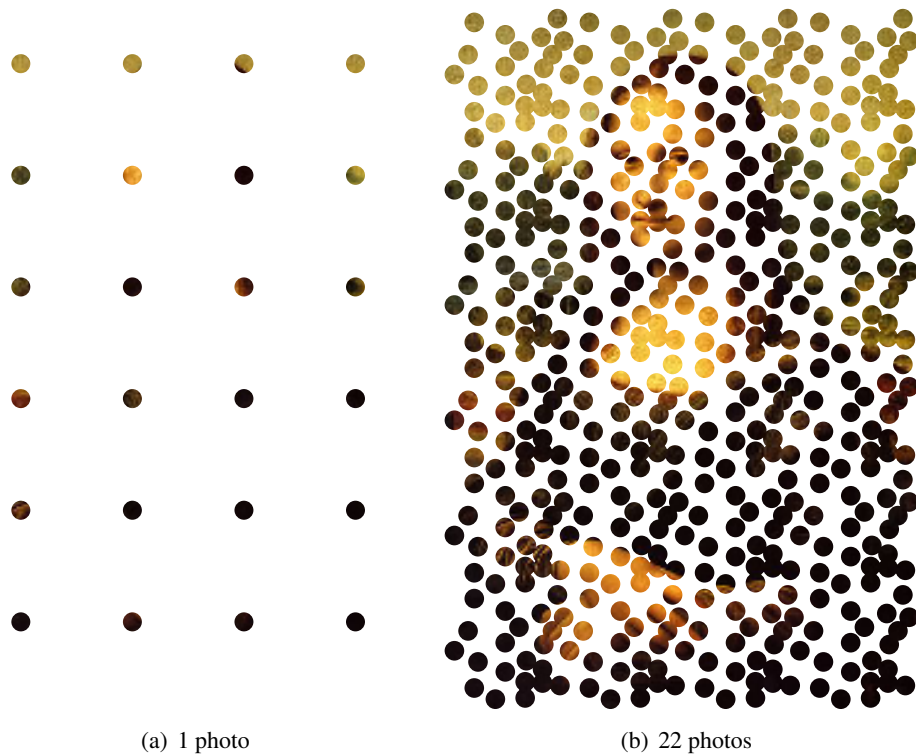


FIGURE 1.2. Échantillonnage par une série de photographies de la figure 1.1 (a).

Une autre façon de produire une image de haute résolution est de faire une hypothèse (dite de régularité) sur le contenu de l'image. Par exemple, on peut supposer qu'entre deux

échantillons l'image varie de manière linéaire. On pourrait aussi supposer que l'image est constante entre les échantillons. Lorsque l'on formule le problème d'interpolation du signal sur une grille plus fine et que l'on ajoute de telles hypothèses, on dit que l'on régularise le problème. Dans tous les cas, la qualité de l'image de haute résolution produite dépendra fortement de la validité de l'hypothèse faite. Sur la figure 1.3, on montre l'importance de faire une régularisation adéquate : le résultat peut être loin de la réalité si l'hypothèse de régularité est erronée. Utiliser une régularisation comporte donc des risques et l'idéal serait de pouvoir s'en passer.

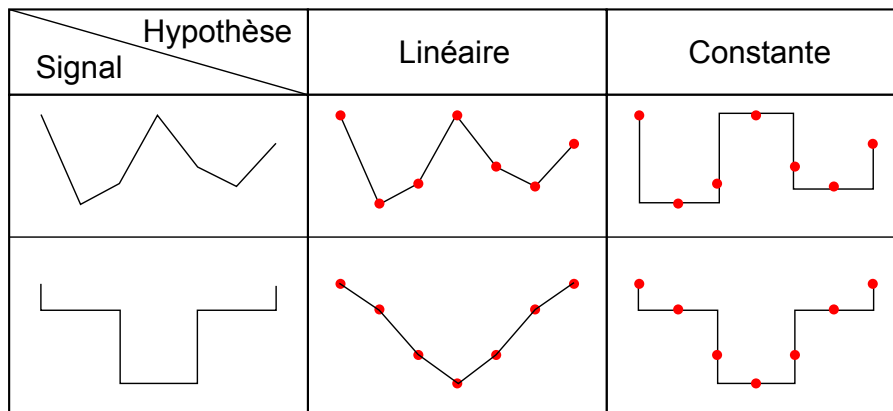


FIGURE 1.3. Régularisation pour l'interpolation. Les signaux sont échantillonnés sur les points rouges. Ils sont ensuite interpolés en utilisant deux modèles de régularité différents.

Le but de cette thèse est de déterminer les conditions qui permettent de retrouver l'image de haute résolution sans régularisation. Lorsqu'une régularisation devient inévitable, on cherche à déterminer quelle force doit avoir l'hypothèse de régularité pour obtenir le résultat le plus fidèle possible.

1.1 Repliement spectral et limites matérielles

Une des principales raisons qui motive la super-résolution est d'éliminer l'effet du repliement spectral dans les images. Ce repliement provoque les effets de crénelages observés sur la figure 1.1 (c). Il est aussi la cause d'effets de moirés sur les structures périodiques des images. Ces effets sont visibles sur la figure 1.4 où la réduction de la résolution provoque l'apparition de structures nouvelles dans l'image. Pour cela, il faut produire une image de la scène à une résolution assez haute pour capturer tous les détails de l'image. Le théorème d'échantillonnage de Shannon garantit qu'il existe une résolution assez fine pour échantillonner correctement l'intensité de la lumière qui arrive sur l'optique d'un appareil photographique, car elle ne comporte pas de fréquences infinies. En effet, le phénomène de diffraction créé par l'ouverture spatiale de l'optique filtre les détails de haute fréquence. Si elle est échantillonnée à cette résolution, l'image ne présentera pas de repliement spectral.

Malheureusement, l'augmentation de la résolution nominale des appareils photo numériques se heurte à une barrière de nature physique, et peine à atteindre la résolution idéale



FIGURE 1.4. Effet de moiré provoqué par une résolution trop faible

donnée par le théorème de Shannon. En effet, la taille des capteurs influe sur le niveau de bruit dans la photographie. Un capteur photosensible peut être vu comme un instrument de comptage de photons pendant le temps de l'ouverture de l'appareil photographique. La valeur du pixel associée à un capteur sera directement liée au nombre de photons qui ont heurté sa surface pendant le temps d'ouverture. Si la taille du capteur diminue, les chances que des photons rencontrent ce capteur diminuent aussi. On peut arriver à des circonstances où aucun photon rencontre le capteur, le niveau de lumière n'est alors pas enregistré. Cela provoque une variabilité dans la mesure de l'intensité moyenne de la lumière par unité de temps. Dans des conditions de lumière faible, le bruit qui résulte de ce phénomène est particulièrement visible. Dans la figure 1.5, on simule l'effet de diviser la largeur du capteur par deux pour augmenter la résolution par deux (figure 1.5). L'image de taille 512×512 ne présente pas de repliement, mais est plus bruitée que celle prise à la résolution inférieure de 256×256 (figure 1.5 (a)).

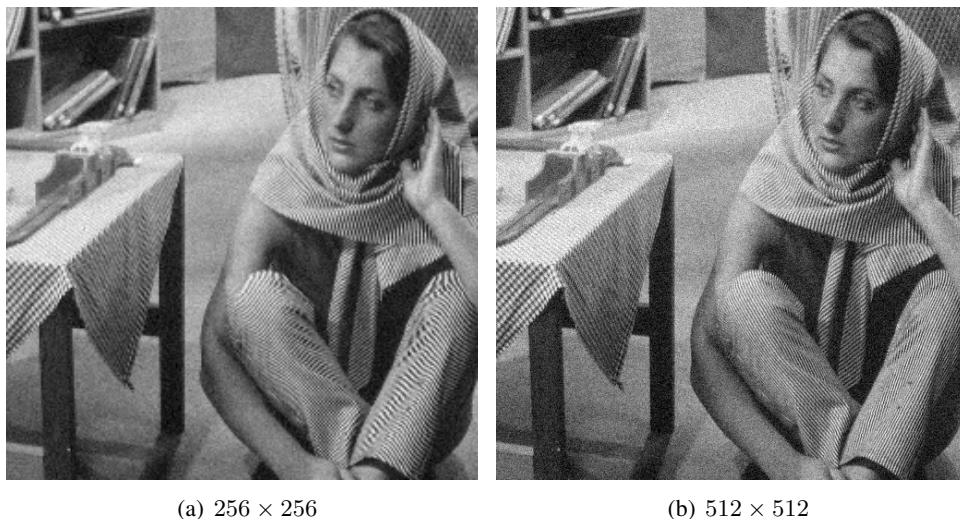


FIGURE 1.5. Compromis résolution / bruit pour les appareils photo numériques.

On ne peut donc pas augmenter indéfiniment la résolution des appareils photo numé-

riques. D'autres méthodes sont nécessaires pour augmenter la résolution finale des photographies, et ainsi limiter le repliement spectral dans les images.

1.2 Problématique

Cela justifie donc le fait de chercher à raffiner l'échantillonnage en prenant plus de photographies. Une des premières questions à laquelle il faut répondre est de déterminer le nombre d'images qui permet de passer de la basse résolution à la haute résolution. Par exemple, avec la scène de la figure 1.4, combien de versions de l'image 1.4 (c) sont nécessaires pour retrouver l'image 1.4 (a) ? Pour des applications réelles, la réponse à cette question devra prendre en compte le bruit d'acquisition présent dans les images. Lorsque le nombre d'images n'est pas suffisant, la régularisation peut être un moyen d'interpoler les échantillons disponibles de manière visuellement satisfaisante. Il faudra cependant l'utiliser avec précaution pour éviter les artefacts causés par une erreur sur l'hypothèse de régularité. De plus, en vue d'une application pratique de la super-résolution, il est important de considérer que des modifications peuvent intervenir dans la scène et de déterminer des algorithmes robustes (résistants) à ces modifications.

1.3 Contributions

On a commencé par déterminer le nombre d'images pour effectuer une super-résolution sans régularisation pour le cas où la scène à prendre en photo ne subit pas de modification, et où le mouvement relatif entre les images est affine (assez général pour décrire de nombreuses situations pratiques). On a montré que, si l'on veut multiplier le nombre d'échantillons d'un facteur M^2 , il est nécessaire et suffisant d'avoir $N = M^2$ images si les images de basse résolution ne sont pas bruitées. On a ensuite montré que dans le cas bruité, il devient possible d'estimer la scène de haute résolution avec une bonne précision lorsque le nombre d'images augmente, des expériences montrent qu'il est possible d'obtenir pratiquement ce nombre d'images. (chapitre 3, publications : [106, 105])

On a utilisé ces résultats dans le cadre d'une super-résolution à haute dynamique où il faut reconstruire toute la plage dynamique de la scène en plus de produire une image de haute résolution. Ces travaux en collaboration avec Cecilia Aguerrebere ont abouti à la détermination de conditions expérimentales qui permettent d'obtenir une image de grande résolution et grande dynamique de bonne qualité. Ces résultats ont été validés par des expériences réelles. (chapitre 4, publications : [104])

Toujours dans le cadre de scènes immobiles, on a étudié le comportement de la régularisation dans le cas où elle devient nécessaire. On a montré que son utilisation pouvait être limitée à certaines zones de l'image, où le bruit produit par le processus de super-résolution est plus fort. On a montré comment déterminer les zones où appliquer le plus de régularisation. (chapitre 5, publications : [106, 105])

On a ensuite étudié le cas où le contenu de la scène change entre les différentes photographies (images contaminées par des données aberrantes ou *outliers*). On a montré que si le pourcentage d'images qui présentent des modifications n'est pas trop grand, on peut

garantir une reconstruction précise de l'image de haute résolution. La super-résolution est robuste aux outliers dans ces conditions. (chapitre 6, publications : [107, 103])

Enfin, on a étudié l'impact d'une classe de méthodes de régularisation (les méthodes parcimonieuses) sur la robustesse du problème de super-résolution. On a montré que si l'hypothèse de régularisation est appropriée pour l'image, et qu'elle est compatible avec la forme des outliers et la super-résolution, alors de telles méthodes peuvent être très robustes. (chapitre 7, publications : [102])

Ces contributions seront détaillées formellement dans la section 2.3.2.

1.4 Publications

Cette thèse a donné lieu aux publications suivantes.

Article de journal

Accepté pour publication :

- [104] **Simultaneous High Dynamic Range and Super-Resolution Imaging Without Regularization**. Yann Traonmilin et Cecilia Aguerrebere. À paraître dans *SIAM Journal on Imaging Sciences*, 2014.

Soumis à révision :

- [105] **On the amount of regularization for super-resolution reconstruction**. Yann Traonmilin, Saïd Ladjal et Andrés Almansa, 2012.
- [102] **Robust Multi-image Processing With Optimal Sparse Regularization**. Yann Traonmilin, Saïd Ladjal et Andrés Almansa, 2014.

Conférences

- [106] **On the amount of regularization for Super-Resolution interpolation**. Yann Traonmilin, Saïd Ladjal et Andrés Almansa. *20th European Signal Processing Conference 2012 (EUSIPCO 2012)*, Bucharest, Romania, August 2012.
- [107] **Outlier removal power of the L1-Norm Super-Resolution**. Yann Traonmilin, Saïd Ladjal et Andrés Almansa. *Scale Space and Variational Methods in Computer Vision, volume 7893 of Lecture Notes in Computer Science*, pages 198209, 2013.
- [103] **Quantification de la robustesse de la super-résolution par minimisation L1**. Yann Traonmilin, Saïd Ladjal et Andrés Almansa. *23ème Colloque Gretsi (Gretsi 2013)*, Brest, France, 2013.
- **How to Define Signal and Noise Sparsity for Multi-image Processing?**. Yann Traonmilin, Saïd Ladjal et Andrés Almansa. *8th International Conference Curves and Surfaces*, Paris, France, 2014.

CHAPITRE 2

MÉTHODES VARIATIONNELLES DE TRAITEMENT DES IMAGES ET SUPER-RÉSOLUTION

Dans ce chapitre, on modélise le problème de l'imagerie de super-résolution et discute des méthodes de reconstruction associées. On s'attache à garder une certaine généralité dans cette discussion. Les détails techniques appropriés seront spécifiés dans chaque chapitre. Cette discussion nous permet de décrire formellement le contenu de nos travaux dans la dernière section de ce chapitre.

2.1 Modèle de super-résolution

2.1.1 Acquisition d'une série d'images

On considère un modèle physique d'acquisition des images par un appareil photo. La lumière est captée par une grille de capteurs photosensibles après avoir traversé une optique. L'énergie incidente sur chaque capteur est ensuite quantifiée pour donner la valeur du pixel correspondant. L'acquisition de l'image est perturbée par du bruit (qui dépend du temps d'exposition ou bien de changements dans l'image...). On peut donc modéliser l'acquisition d'une image numérique w_i par :

$$w_i = f(SFQ_i u_0 + n_i), \quad (2.1)$$

où u_0 est l'irradiance de la scène dans un repère donné, n_i est le bruit qui perturbe l'acquisition, Q_i est la déformation appliquée à cette scène avant d'atteindre l'optique. F est le flou de l'appareil. S est l'échantillonnage sur la grille de capteurs à un pas δ_{BR} , et f est la fonction qui regroupe l'effet de la transformation du flux de photons en information quantifiée. On appelle la grille d'acquisition, grille basse résolution (BR). Tous ces opérateurs sont linéaires sauf f . Dans nos travaux, le bruit de quantification est négligé, ce qui revient à l'inclure dans le bruit d'acquisition. La saturation des capteurs est considérée seulement dans le chapitre 4. Lorsque l'on considère l'acquisition de N images, on considère l'opérateur linéaire A agissant sur une scène d'irradiance quelconque u :

$$\begin{aligned} A : \\ u &\rightarrow (SFQ_i u)_{i=1,N}. \end{aligned} \quad (2.2)$$

On arrive finalement au modèle d'acquisition :

$$w = Au_0 + n \quad (2.3)$$

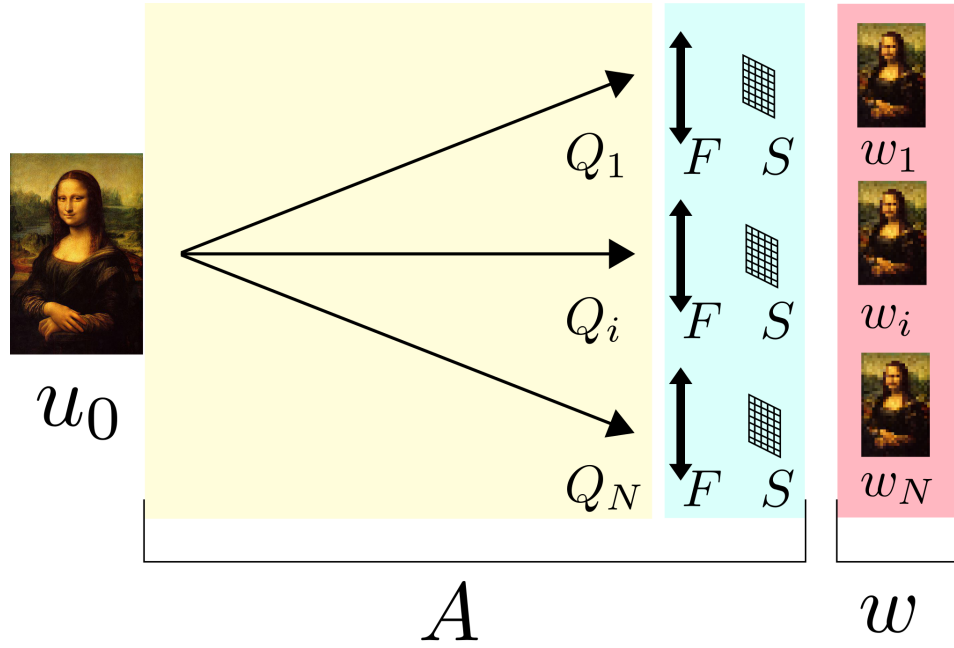


FIGURE 2.1. Acquisition d'une série d'images.

où w est la concaténation des images basse résolution et n la concaténation des bruits n_i . On appelle u_0 , l'image haute-résolution (HR). On représente cette chaîne d'acquisition figure 2.1.

Avec ces notations, on est en mesure de donner une définition du problème d'imagerie par super-résolution.

Définition 2.1.1. *Le problème d'imagerie par super-résolution consiste à retrouver u_0 à partir de $w = Au_0 + n$. On appelle aussi ce processus reconstruction par super-résolution (reconstruction SR).*

En pratique, les Q_i et F sont inconnus. On distingue un sous problème de super-résolution, dans le cas où A est connu :

Définition 2.1.2. *Supposons que l'opérateur A est connu. On appelle le problème d'imagerie SR, interpolation par super-résolution (interpolation SR).*

En effet, si A est connu, le problème de SR revient à effectuer un rééchantillonnage d'une grille irrégulière vers une grille régulière. Ce modèle d'acquisition multi-image permet de modéliser une grande gamme de problèmes de traitement d'images :

- Pour $M = 1$, $F = I$, on parle de problème de débruitage multi-image.
- Si $N = 1$ et $M = 1$, on parle d'estimation de flou et de déconvolution.
- Si $N = 1$, $F = I$, il s'agit d'un problème d'interpolation classique (aussi appelé super-résolution mono-image).

2.1.2 Description des paramètres du modèle

Le problème principal de modélisation auquel on est confronté est de déterminer les espaces dans lesquels vivent les objets u_0 et w qui modélisent l'irradiance et les images numériques. Idéalement, pour un modèle continu de l'irradiance, u_0 devrait être une fonction de $\mathbb{R}^2 \rightarrow \mathbb{R}$ et chaque w_i un vecteur de $\mathbb{R}^{l \times l}$ où l est la taille de l'image. Certains travaux s'attachent à résoudre ce problème d'interpolation général [20, 33] en une dimension (1D). Lorsque l'on implémente un système de super-résolution pratiquement, on est restreint à la résolution d'une approximation discrète et finie du problème.

Discretisation : L'effet d'ouverture spatiale de l'appareil photo agit comme un passe-bas parfait. On peut donc supposer que l'irradiance échantillonnée par l'appareil est à bande limitée (une partie de son contenu fréquentiel a disparu au-delà de la fréquence de coupure du filtre). Il existe donc un pas d'échantillonnage haute résolution δ_{HR} tel que u_0 est bien échantillonnée. Sans perte de généralité, on peut donc considérer que l'irradiance u_0 est une fonction de $\mathbb{Z}^2 \rightarrow \mathbb{R}$. On appelle le rapport $M = \delta_{BR}/\delta_{HR}$ le facteur de super-résolution. Dans cette thèse, M est entier, ce qui ne constitue pas une limitation sévère car on peut toujours remplacer δ_{BR}/δ_{HR} par l'entier supérieur et avoir une reconstruction légèrement redondante.

Restriction spatiale : Comme vu précédemment, les images BR sont de taille finie. On ne peut donc pas espérer retrouver l'image u_0 de taille infinie. En pratique, on cherchera donc à retrouver seulement $l' \times l'$ pixels de u_0 , où $l' = M \times l$. C'est ici que la modélisation discrète finie atteint une limite. En effet on ne peut pas modéliser les mouvements (Q_i) sur des images discrètes infinies par des mouvements sur des images discrètes finies car obtenir un pixel de l'image translatée nécessite la multiplication de l'image par un sinus cardinal de taille infinie. Cependant, cette opération peut être effectuée en modulant la transformée de Fourier à temps discret. Dans ce travail, on suppose que l'on peut appliquer, de manière similaire, le mouvement Q_i par interpolation par sinus cardinal discret, et donc par modulation de la transformée de Fourier discrète et finie (TFD). Ce qui revient à appliquer la transformation sur une version périodisée (infinie) de l'image. Cette opération est possible s'il n'y a pas de discontinuité créée lors de cette transformation de l'image périodisée [18, 54, 73]. Avec une apodisation adéquate de l'image, on limite l'impact en fréquence de la restriction spatiale. On peut de plus montrer que l'image fenêtrée est une bonne approximation de l'image infinie si la fenêtre est assez grande [30]. On se restreint donc au problème approché où $u_0 \in \mathbb{C}^{l' \times l'}$ et $w_i \in \mathbb{C}^{l \times l}$.

Types de bruit : Le bruit d'acquisition n joue un rôle important dans la difficulté de la SR. Il est la somme de toutes les perturbations par rapport au modèle que l'on vient de présenter. On peut distinguer deux types de bruits différents :

1. Le bruit de capture : il dépend directement de la précision avec laquelle les capteurs captent l'intensité de lumière moyenne par unité de temps. Il est souvent modélisé comme une variable aléatoire, de manière générique ou tenant compte précisément de la physique des capteurs (Section 4). En général, on suppose que ce type de bruit est d'énergie finie. Il est souvent modélisé comme une variable aléatoire gaussienne ou poissonnienne.

2. Le bruit de type corruption de données (outliers) : Tous les changements qui peuvent perturber l'image tels que les occultations, l'apparition de reflets, etc... Dans ce cas, le bruit peut avoir une énergie arbitrairement grande.

Mouvement : Différentes complexités de mouvement ont été étudiées dans la littérature. Par ordre de difficulté croissante : des mouvements translationnels ont d'abord été considérés [45], puis des compositions de translations et rotations [58, 110, 57], des mouvements affines [116] et des homographies [92]. Des mouvements génériques décrits en utilisant un flot optique ont aussi été considérés [46]. Dans cette thèse, on considère que le mouvement relatif entre les images est affine. Ce choix est un compromis entre simplicité et capacité de description réelle du mouvement. Théoriquement, le déplacement observé entre deux acquisitions d'un plan est une homographie (fonction non linéaire des coordonnées) [55]. Cependant lorsque les mouvements ne sont pas trop grands, cette homographie est bien approximée localement par un mouvement affine. On décrit en détail le modèle que l'on utilise dans la section 3.1.1 du chapitre 3.

Flou d'appareil : Le flou d'appareil est la réponse impulsionnelle (ou l'effet de la fonction de transfert) de la partie optique et de la partie photosensible de l'appareil. Dans beaucoup de travaux, l'estimation du flou est découplée de l'estimation de l'image HR. C'est-à-dire que l'image HR floutée est estimée, le flou étant estimé et enlevé ensuite [45]. D'autres travaux cherchent à estimer le noyau de flou en même temps que les mouvements et l'image HR. Un problème récurrent de la littérature sur la super-résolution est la présentation systématique de résultats où la déconvolution est appliquée. Il est alors difficile de déterminer si l'amélioration de la résolution du résultat est due au rehaussement des hautes fréquences de la déconvolution ou à un dépliement du repliement. Cet effet est illustré figure 2.2. On compare l'image HR avec flou (résultat d'une super-résolution où l'on n'a pas inversé le flou) et le résultat d'une mauvaise interpolation suivie d'une déconvolution (déconvolution exacte dans le domaine de Fourier). L'image déconvoluée semble plus nette, mais certaines parties sont mal interpolées. Si on n'avait pas l'image de référence en comparaison, il serait difficile de conclure que le résultat d'une SR parfaite, avec ou sans déconvolution, est plus fidèle. Si on montre un résultat avec la déconvolution appliquée, il est donc difficile de déterminer quelle part de l'augmentation de la netteté est due à la déconvolution, et quelle part est due à l'interpolation. Dans ce travail, l'estimation et la déconvolution du flou n'ont pas été étudiées. On justifie plus précisément comment effectuer cette simplification dans la section 3.1.1.

2.2 Problèmes inverses d'imagerie

La formulation : $w = Au_0 + n$ mène aux problèmes inverses (linéaire lorsque A est connu et linéaire) de traitement des signaux. On cherche à retrouver un signal u_0 à partir de son observation w . Les méthodes dites variationnelles consistent à essayer de retrouver u_0 en minimisant la somme de deux fonctionnelles (en toute généralité u_0 est une fonction, même si en pratique elle est discrétisée) :

$$\tilde{u} = \operatorname{argmin}_u D(u, w) + \lambda R(u) \quad (2.4)$$

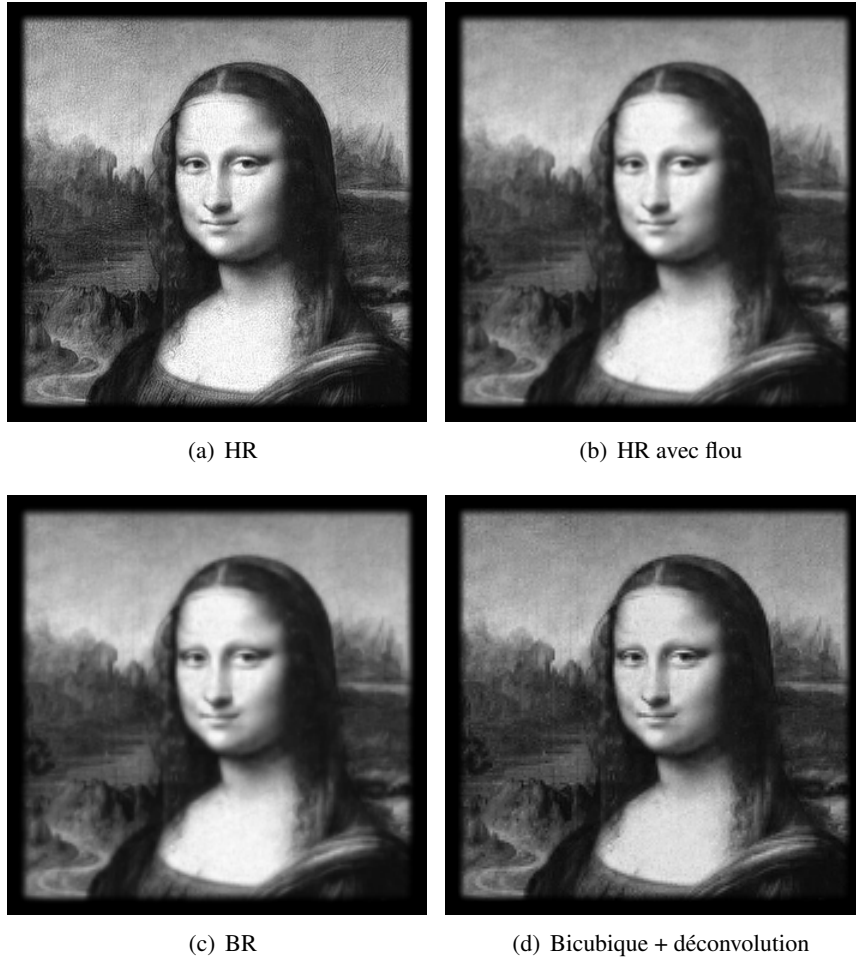


FIGURE 2.2. Effet de la déconvolution sur une image mal interpolée (a) l'image HR, (b) l'image HR avec flou, (c) image BR (d) résultat d'une interpolation bicubique suivie de la déconvolution.

où D est une fonction de distance qui quantifie la distance d'une solution potentielle aux observations que l'on appelle attache aux données. La fonction R est une fonction de régularisation qui quantifie l'éloignement par rapport à un modèle de régularité pour u_0 . De cette manière, on agrège la connaissance que l'on a de la scène (l'attache aux données) et celle sur le modèle que l'on en fait. Au final, l'objectif est le suivant : pour un ensemble E qui permet de modéliser la régularité des inconnues $u_0 \in E$, la minimisation (2.4) doit donner $\tilde{u} = u_0$. Plus l'ensemble E est grand, plus les hypothèses sur u_0 sont réduites. Pour le cas du traitement des images numériques, l'idéal est que E contiennent toutes les images naturelles.

Une grande quantité de méthodes reposent sur la minimisation d'une distance de norme L^p pour D .

$$\tilde{u} = \operatorname{argmin}_u \|Au - w\|_p^p + \lambda R(u) \quad (2.5)$$

Les candidats sont nombreux pour R et dépendent des hypothèses que l'on fait sur les

images. De la même manière, il s'agit de minimiser une norme L^q :

$$\tilde{u} = \operatorname{argmin}_u \|Au - w\|_p^p + \lambda \|\Psi u\|_q^q \quad (2.6)$$

où Ψ est l'opérateur (souvent linéaire) de régularisation. On peut ainsi retrouver plusieurs algorithmes classiques de résolution de problèmes inverses :

- Si $p = 2$, on parle de solution aux moindres carrés.
- Si $p = 1$, on parle de régression robuste (LASSO dans le monde des statistiques).
- Si $q = 2$, on parle de régularisation de Tychonoff.
- Si $q = 1$ et Ψ est le gradient, c'est une régularisation par variation totale (TV pour « Total Variation ») anisotrope.
- Si $q = 1, 0$ et Ψ est une projection dans un dictionnaire redondant, on parle de régularisation parcimonieuse.

Remarque 2.2.1. *En toute généralité, les inconnues sont ici des fonctions, d'où la notation L^p pour désigner les normes. Dans cette thèse, on s'est intéressé seulement à une version discrétisée du problème de super-résolution, on aurait donc pu utiliser la notation ℓ^p habituellement utilisée pour les suites.*

La question principale lors de l'étude de problèmes inverses est de savoir s'il est possible de trouver l'inconnue. Tout d'abord, si aucune hypothèse de régularité n'est disponible, est-on capable de retrouver u_0 ?

Définition 2.2.1 (Inversibilité). *Le problème (2.4) est dit inversible si pour $\lambda = 0$, \tilde{u} existe et est unique.*

Définition 2.2.2 (Problème mal posé). *Le problème de l'équation (2.4) est dit mal posé si pour $\lambda = 0$, \tilde{u} n'est pas unique.*

La régularisation semble inévitable lorsque le problème est mal posé, c'est-à-dire lorsque la minimisation pour $\lambda = 0$ (attache aux données seule) n'admet pas une solution unique ou a une solution très éloignée de u_0 . C'est le cas lorsque A n'est pas injective. Le problème d'interpolation mono-image en est un exemple classique.

Faire une hypothèse de régularité sur l'inconnue permet de lever l'ambiguïté lorsque l'opérateur n'est pas injectif, ou de réduire le bruit si ce dernier ne suit pas cette hypothèse de régularité.

2.2.1 Problèmes convexes et non convexes

On rappelle qu'une fonction $f : \mathbb{R}^p \rightarrow \mathbb{R}$ avec $f(x) = f(x)$ est convexe si pour tout $x, y \in \mathbb{R}^p$ et $\lambda \in [0, 1]$, on a

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (2.7)$$

On cherche souvent à savoir si la minimisation de l'équation (2.4), est la minimisation d'une fonction convexe ou non. En effet, si une fonction est convexe, tout minimum local est minimum global. Il est possible de minimiser une telle fonction grâce à des algorithmes rapides (particulièrement si elle est différentiable). Si la fonction à minimiser n'est pas convexe, elle peut alors avoir des minimums locaux, qui rendent la minimisation coûteuse en temps.

Par exemple, le problème de minimisation de normes L^p, L^q de l'équation (2.6) est convexe si $p, q \geq 1$ et non convexe en général si $p < 1$ ou $q < 1$.

2.2.2 Le cas de la super-résolution

En général, dans le cas de la super-résolution, on ne connaît pas l'opérateur A . On cherche donc à minimiser

$$\tilde{u} = \operatorname{argmin}_{A,u} \|Au - w\|_p^p + \lambda \|\Psi u\|_q^q. \quad (2.8)$$

L'opérateur A est tout de même restreint à un certain ensemble (mouvements affines). Ce problème n'est pas convexe et est en général considéré comme mal posé et mal conditionné pour $\lambda = 0$. C'est-à-dire que l'on se repose sur un modèle théorique de l'image pour apporter l'information manquante à la résolution du problème. Dans le cas de la super-résolution, le résultat d'une telle minimisation n'est donc pas en général, l'image digitale qui aurait été capturée si le pas d'échantillonnage avait été M fois plus petit.

L'estimation de A recouvre l'estimation du flou de l'appareil et des paramètres de mouvement de chaque image. Des avancées sur la question de l'estimation du flou [36] montrent que cette partie du problème peut être bien posée et séparée du reste du problème de SR. Lorsque l'on fixe A , le problème devient linéaire. Pour des cas particuliers, on sait déjà que ce sous-problème est bien posé, et qu'il n'est pas nécessaire de régulariser. Par exemple, si l'échantillonnage qui résulte de la fusion des grilles d'échantillonnage de basse résolution est régulier.

La question se pose donc de savoir quelles sont les conditions les plus générales possibles qui permettent de s'abstenir de régulariser, et ainsi obtenir le résultat le plus fidèle possible à la scène réelle. On déterminera ainsi quand il est utile d'utiliser une large classe de méthodes de super-résolution.

Méthodes $L^p - L^q$ de l'état de l'art : Beaucoup de méthodes de l'état de l'art peuvent être formulées sous cette forme variationnelle. Elles sont décrites extensivement dans les revues de l'état de l'art de la littérature [80, 43, 70, 101]. On trouve fréquemment, dans la littérature, une interprétation probabiliste équivalente à la formulation variationnelle (recherche d'un maximum a posteriori utilisant un a priori sur les données). On peut trouver une description de cette équivalence dans le chapitre 7 de [17]. Par exemple, [53] effectue un maximum a posteriori qui revient à une minimisation $L^2 - L^2$, avec Ψ l'inverse d'une estimation de la matrice de covariance de l'image HR. Dans la suite, on utilisera seulement le vocabulaire variationnel pour décrire les méthodes de manière unifiée. Dans [39], des algorithmes d'estimation adaptative (*steepest descent* (méthode de descente de gradient) tronquée, *recursive least squares* (moindres carrés récursifs)) pour la minimisation L^2 sont comparés. [40] propose un algorithme pour effectuer la super-résolution L^2 sans régularisation. [85] propose une minimisation $L^2 - L^2$ où Ψ est le gradient discret. [76] étudie les propriétés des méthodes $L^2 - L^2$ ainsi que la convergence d'algorithmes associés. [45] propose un modèle d'attache aux données L^1 avec régularisation par variation totale anisotrope (appelée bilatérale), ce qui est une minimisation $L^1 - L^1$. L'avantage principal de cette méthode est sa robustesse aux données aberrantes. [65, 68] proposent d'effectuer la minimisation de l'attache aux données L^2 avec régularisation de la variation totale pour la super-résolution mono-image. [115] propose de pondérer la variation totale pour qu'elle soit moins dépendante au paramètre de régularisation. [10] utilise une minimisation $L^2 - TV$ (proche de $L^2 - L^1$) pour produire l'image HR dans son algorithme général de reconstruction SR.

Autres méthodes : On peut distinguer d'autres types de méthodes. Certaines suivent aussi un modèle attache aux données plus régularisation, mais n'utilisent plus de simples normes L^q . Par exemple, [94] utilise un a priori de Huber-Markov (la fonction R est une norme L^1 lissée autour de 0, ce qui revient aussi à prendre Ψ non-linéaire). Dans [81], une approximation polynomiale locale du problème est utilisée. [112] présente un algorithme EM (Expectation - Maximization) pour le problème de reconstruction SR. [84] généralise l'algorithme des Non Local Means (débruitage par auto-similarité) au problème de la super-résolution, et en déduit un algorithme où le mouvement n'est pas explicitement estimé. Dans [98], la méthode de Steering Kernel Regression, où on estime localement la covariance de l'image, est appliqué à la super-résolution. Cette méthode n'utilise pas non plus d'estimation explicite du mouvement. [100] donne une méthode basée sur les chaînes de Markov Monte Carlo.

2.2.3 Modèles de régularité pour les images

À partir du moment où un problème inverse d'imagerie est non inversible, il est nécessaire d'avoir une autre source d'information sur l'inconnue pour pouvoir espérer produire une solution valide. Cette source d'information prend la forme d'un modèle de régularité sur l'image.

Comme on l'a vu dans la section précédente, la forme de l'équation (2.8) permet de modéliser une large gamme (non exhaustive) de modèles. La plupart de ces modèles ont été appliqués à la super-résolution. On peut distinguer deux grandes classes de régularisation. Les méthodes qui font une hypothèse sur leur caractère lisse et les méthodes à base de dictionnaires.

Pour mesurer les variations d'une image, le gradient est le plus souvent utilisé. Lorsque Ψ est l'opérateur de gradient et que l'on minimise $\|\Psi u\|_2^2$, on parle de régularisation laplacienne. Lorsque l'on minimise $\|\Psi u\|_1$, il s'agit de la minimisation de la variation totale anisotrope. En effet, la quantité $\|\Psi u\|_1$ est la somme des variations absolue d'une image. On peut mentionner aussi l'hypothèse de $\|\Psi u\|_0$ minimal que l'on appelle hypothèse de gradient parcimonieux.

Dans les méthodes à base de dictionnaires [38], Ψ est une projection dans un dictionnaire. Ce dictionnaire est, soit construit à partir de l'image, soit fourni par ailleurs. De la même manière que pour le gradient, on cherche à minimiser l'énergie de la représentation dans cette famille d'atomes (norme L^2), la somme des valeurs absolues des coefficients de la décomposition (norme L^1) ou maximiser la parcimonie (norme L^0).

2.3 Description formelle de la problématique

2.3.1 Super-résolution bien informée

Dans cette thèse, on s'est posé la question suivante : dans quel cas est-il possible d'éviter de faire une hypothèse de régularité sur l'image ? Dans le cas où l'on est obligé de la faire, comment minimiser cette hypothèse ?

Le principal facteur qui va influencer sur cette question est le nombre d'images disponibles pour effectuer cette super-résolution, c'est pourquoi on parlera de quantité d'information suffisante pour effectuer la super-résolution. Cette quantité d'information (le nombre d'images en particulier) est d'ailleurs un des seuls paramètres de l'acquisition que l'on peut maîtriser. On introduit cette notion pour rassembler les notions déjà usuelles de système

bien conditionné (et donc inversibles), de quantité de bruit de reconstruction (aussi appelé stabilité de l'inversion).

Définition 2.3.1. *Étant donné un ensemble de paramètres A, n , on dit que la super-résolution est parfaitement informée pour ces paramètres et pour les images appartenant à l'ensemble E , s'il existe un algorithme \mathcal{A} prenant w en entrée tel que pour tout $u_0 \in E$, $\mathcal{A}(w) = u_0$.*

En pratique, on ne pourra jamais obtenir une super-résolution parfaitement informée, car le bruit qui contamine les données l'en empêche. On considère que la résolution est bien informée si l'on peut approcher l'image idéale avec une certaine précision.

Définition 2.3.2. *Étant donné un ensemble de paramètres A, n , on dit que la super-résolution est bien informée avec une précision ϵ pour ces paramètres et pour les images appartenant à l'ensemble E , s'il existe un algorithme \mathcal{A} prenant w en entrée, tel que pour tout $u_0 \in E$,*

$$\|\mathcal{A}(w) - u_0\| \leq \epsilon \quad (2.9)$$

Dans ces deux définitions, lorsque l'on ne précise pas E (pas de modèle de régularité sur l'image), il s'agira de l'ensemble des images bien échantillonnées au pas HR.

La question de cette thèse devient donc : pour quels A, n la super-résolution est elle bien informée avec une bonne précision ? Que faire lorsque cette précision n'est pas assez grande ?

2.3.2 Plan de la thèse revisité

Pour cette étude, on s'est concentré sur l'étude de deux schémas de minimisation d'attache aux données : les minimisations L^2 et L^1 . Comme la manière de régulariser revient à choisir un modèle de données, on s'abstient de juger la qualité d'un modèle plutôt qu'un autre, puisque celle-ci dépendra sûrement de l'application. On s'est concentré sur un type de régularisation, la minimisation de la variation totale, car elle est bien installée, avec ses différentes versions, dans le paysage du traitement des images. Sa formulation relativement simple dans le cadre décrit dans la section précédente a également joué un rôle. Il faut garder à l'esprit que la partie de ce travail impliquant de la régularisation (et donc la variation totale) est une base pour l'étude de schémas de régularisation différents, suivant le modèle de l'équation (2.6).

Grâce à la description du modèle et des méthodes de SR que l'on vient d'effectuer, on peut formellement décrire le contenu de chaque partie de cette thèse.

Dans le chapitre 3, on montre que l'interpolation SR affine non bruitée est presque sûrement parfaitement informée si le nombre d'images N est plus grand que M^2 . Lorsque le problème est bruité, pour des mouvements aléatoires, on montre que lorsque le nombre d'images augmente, l'interpolation SR affine devient bien informée avec une certaine précision (qui tend vers 0 dans le cas de bruit Gaussien) avec grande probabilité. La vitesse de convergence peut être calculée grâce à l'étude asymptotique du conditionnement du problème. Lorsque l'on considère le problème de reconstruction SR, on montre que si une estimation initiale des paramètres de mouvements est disponible, la reconstruction devient bien informée lorsque le nombre d'image augmente. On utilise ces résultats dans le chapitre 4 pour déterminer des paramètres pratiques d'acquisition d'images en vue d'effectuer

une super-résolution à haute dynamique.

Dans le chapitre 5, lorsque le nombre d'images devient proche du cas critique M^2 , la précision de la SR n'est pas bonne. On étudie précisément de quelle manière le bruit est généré. Dans le cas de petits mouvements (simulant un appareil photo tenu à la main), on observe une variabilité spatiale du bruit de reconstruction par minimisation L^2 . On montre qu'il est possible de prédire cette variabilité de manière rapide. On en déduit une méthode de régularisation locale, qui vise à reconstruire au maximum les détails de haute fréquence de l'image.

Dans le chapitre 6, on étudie dans quel cas l'interpolation SR L^1 est parfaitement informée, lorsque celle-ci est contaminée par du bruit de type outlier. On trouve que si le ratio d'images contaminées sur le nombre d'images total est inférieur à une constante (dépendant de la taille des images), alors cette interpolation est parfaitement informée.

Dans le chapitre 7, on se pose la question suivante, est ce qu'ajouter une hypothèse de parcimonie sur u_0 augmente le nombre d'images pour lequel l'interpolation SR L^1 est parfaitement informée ? Suivant la structure spatiale des outliers et le domaine dans lequel le signal est parcimonieux, la réponse peut varier. Pour des observations aléatoires, un compromis entre la parcimonie du signal et du bruit doit être fait. Dans le cas du traitement multi-image, la régularisation TV n'apporte rien pour des outliers sans structure spatiale prescrite. En revanche, en imposant une structure sur les outliers, on peut déterminer des conditions pour lesquelles l'interpolation SR est bien informée, ainsi que le paramètre de régularisation nécessaire pour l'effectuer dans des cas particuliers.

CHAPITRE 3

NOMBRE D'IMAGES POUR UNE SUPER-RÉSOLUTION BIEN INFORMÉE

Dans ce chapitre, on cherche à déterminer le nombre d'images pour lequel le problème de super-résolution est bien informé. En premier lieu, on étudie le cas de l'interpolation non bruitée, ce qui revient à étudier l'inversibilité du système linéaire associé au problème d'interpolation de super-résolution. On s'intéresse ensuite à la super-résolution avec bruit additif d'énergie finie. On montre que l'interpolation devient bien informée avec grande probabilité lorsque le nombre d'images augmente. L'estimation des paramètres bénéficie aussi de cette augmentation.

3.1 Inversibilité de l'interpolation SR affine

Dans le cas de l'interpolation, la SR peut être vue comme un problème de rééchantillonnage régulier à partir de données irrégulières, et pourrait donc être résolue à l'aide de techniques associées [51, 41, 9, 6]. Cependant, avec la configuration particulière de l'échantillonnage de la super-résolution, on peut utiliser le fait que chaque image est enregistrée sur une grille régulière, pour élaborer des résultats et méthodes de SR dédiés. Par exemple, le cas de la SR translationnelle peut être vu comme un problème d'échantillonnage multicanal. Ahuja et Bose [4] ont étendu à la 2D le résultat 1D de Papoulis [79], et ont montré que si l'on effectue une super-résolution translationnelle 2D de facteur M , seulement M^2 images BR sont nécessaires pour retrouver l'image HR parfaite dans le cas non bruité. Dans ce cas, on ne fait pas de supposition sur le contenu de l'image HR (en dehors du fait qu'elle est à bande limitée), et la super-résolution est bien informée.

Dans cette section, on définit précisément les hypothèses qui nous permettront d'étudier l'inversibilité dans le cas de mouvements affines. On donne une condition suffisante d'inversibilité de la SR 2D affine. Cela étend le travail d'Ahuja et Bose [4] qui est restreint aux mouvements translationnels. Ce résultat montre que le nombre d'échantillons suffisant est le même que pour l'échantillonnage aléatoire des polynômes trigonométriques décrit par Bass & Gröchenig [13].

3.1.1 Hypothèses

On utilise ici le modèle discret fini pour la super-résolution affine défini comme suit :

$$\begin{aligned} A : (\mathbb{C}^{Ml \times Ml}) &\rightarrow (\mathbb{C}^{l \times l})^N \\ u &\rightarrow Au = w = (SQ_i u)_{i=1, N} \end{aligned} \tag{3.1}$$

où N est le nombre d'images BR, M le facteur de super-résolution, $l \times l$ est la taille des images BR, u est une image HR, l'opérateur S est le sous-échantillonnage d'un facteur M

et les Q_i sont les transformations affines subies par chaque image acquise. On appelle q_i la transformation affine correspondante sur les coordonnées de la grille d'échantillonnage. C'est-à-dire

$$(Q_i u)(\mathbf{x}) = u(q_i(\mathbf{x})). \quad (3.2)$$

On décompose q_i en une partie linéaire et une partie translationnelle : $q_i \mathbf{x} = l_i \mathbf{x} + t_i$. Les l_i sont des matrices 2×2 , t_i sont des vecteurs de \mathbb{R}^2 , le mouvement de chaque image est donc décrit par 6 paramètres.

Des suppositions sont nécessaires pour simplifier les développements théoriques de ce chapitre :

1. Les images HR u sont à bande limitée.
2. La première image BR est l'image de référence ($Q_1 = Id$).
3. Les transformations affines sur les coordonnées (q_i) sont inversibles.
4. Les affinités ne génèrent pas de repliement spectral sur la grille BR.

La condition 1) est toujours remplie, car la diffraction causée par l'ouverture physique de l'appareil photo agit comme un filtre passe-bas parfait. En pratique, avec des appareils réels, un facteur de SR de 2 à 4 peut être considéré [70]. La condition 2) est technique. La condition 3) est toujours remplie lorsque les affinités ne sont pas dégénérées (leurs parties linéaires sont inversibles).

La condition 4) a besoin de plus d'explications. D'après le modèle continu de formation de l'image décrit chapitre 2, les images BR sont générées par $w_i = S_{HR} Q_i u_0$, où S_{HR} est l'échantillonnage de haute résolution. Dans un algorithme de SR pratique, on cherche à estimer $S_{HR} u_0$, c'est-à-dire commuter l'échantillonnage HR et le mouvement. Cette opération est possible si u est bien échantillonné avant et après le mouvement. On montre ici que c'est équivalent à un choix adéquat du pas d'échantillonnage HR.

Proposition 3.1.1. *Soit u_0 une image continue à bande limitée. Il existe un facteur de zoom M et un opérateur de sous-échantillonnage S_{HR} tel que $S_{HR} Q_i u_0 = Q_i S_{HR} u_0$ et $S Q_i u_0 = S S_{HR} Q_i u_0$. Le facteur M est le ratio des pas d'échantillonnage S et S_{HR} .*

Démonstration. Soit D le support du spectre de u_0 . Soit D_i le support du spectre de $Q_i u_0$. Alors $D_i = l_i^{-T}(D)$ (d'après les propriétés de la transformée de Fourier, avec $l_i^{-T} = (l_i^{-1})^T$). Alors $S_{HR} u_0$ est bien échantillonné après chaque mouvement si $\bigcup_i D_i \subset D_{HR}$, où D_{HR} est la cellule réciproque de S_{HR} (le domaine spectral où le théorème de Shannon est valide au pas d'échantillonnage HR [7]). On choisit le pas d'échantillonnage HR le plus grand possible (l'opérateur S_{HR}) tel que $\bigcup_i D_i \subset D_{HR}$. La configuration de ces domaines fréquentiels est montrée dans la Fig. 3.1. \square

En d'autres termes, pour un signal à bande limitée donné, on peut toujours définir un pas d'échantillonnage HR qui ne provoque pas de repliement de l'image continue u_0 après mouvement. La force de cette hypothèse dépend de l'amplitude des transformations affines. Plus elles sont petites (ex. : mouvement induit par un appareil photo tenu à la main), plus l'hypothèse est faible.

Remarque 3.1.1. *Pour cette proposition, on se place dans le cas non fenêtré. Le fait de traiter des images finies introduit des effets de bords que l'on néglige ici (voir Chapitre 2).*

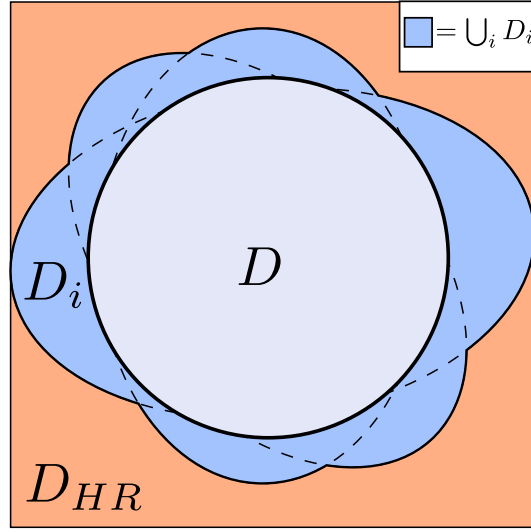


FIGURE 3.1. Configuration de l'échantillonnage dans le domaine de Fourier. Représentation du spectre du signal (D et $\bigcup_i D_i$) et de la limite spectrale de Nyquist de la grille HR (D_{HR})

À propos du flou d'appareil

Normalement, un effet de flou perturbe le signal (noté F avant le sous-échantillonnage) :

$$w_i = SFQ_i u \quad (3.3)$$

Dans ce chapitre, on n'essaie pas d'enlever ce flou : on suppose que le flou est identique pour toutes les images BR et que l'on peut commuter flou et mouvement. On cherche donc à estimer Fu . Une condition suffisante pour le faire en inversant l'opérateur A (qui ne comporte pas le flou) est : $FQ_i = Q_i F$. C'est une hypothèse souvent faite dans la littérature [45, 116]. Cette supposition est vraie pour les mouvements purement translationnels et pour les rotations tant que le noyau de flou est invariant par rotation.

Cette hypothèse peut se justifier pour des noyaux de flou plus généraux dans le cas de petits mouvements. La différence entre l'estimation de l'image HR par minimisation L^2 , que l'on appelle \tilde{u} (détaillée dans la section 3.2.1), lorsque l'on n'inclue pas le flou dans A , et l'image HR filtrée par ce flou Fu_0 est :

$$\begin{aligned} \tilde{u} - Fu_0 &= (A^H A)^{-1} \left(\sum_i Q_i S^H w_i - (A^H A) Fu \right) \\ &= (A^H A)^{-1} \left(\sum_i Q_i S^H S(FQ_i - Q_i F) u_0 \right) \end{aligned} \quad (3.4)$$

L'hypothèse est valide si l'énergie de $S(FQ_i - Q_i F)u_0$ est plus petite que l'énergie du bruit d'acquisition. On montre une expérience de reconstruction dans la figure 3.2 : en négligeant le flou dans le modèle, on reconstruit une version approchée de l'image HR floue.

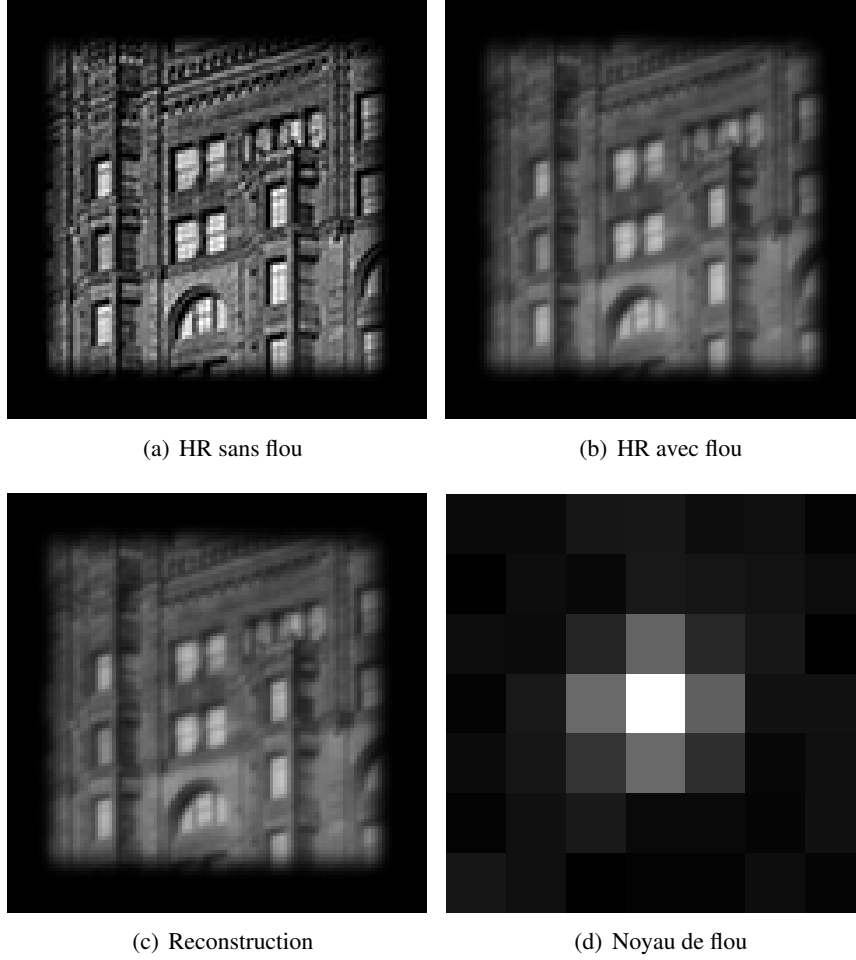


FIGURE 3.2. Validité de l'hypothèse de commutativité flou-mouvement. (a) image HR (b) image HR floutée (c) image reconstruite avec 16 images BR et $M = 2$ en négligeant le flou. PSNR calculé par rapport à (b) est de 46.7 (d) noyau de flou.

3.1.2 Théorème d'inversibilité

Le problème d'interpolation SR sans bruit est un système d'équations linéaires. Suivant le nombre d'images ce système est sous-déterminé ou surdéterminé. La taille du système d'inconnue u (avec $w = Au_0$)

$$Au = w, \quad (3.5)$$

est $Nl^2 \times M^2l^2$. Il est donc nécessaire que $N \geq M^2$ pour que le problème soit inversible. Comme $w = Au_0$, l'inversibilité de ce système est équivalente à l'injectivité de A . Si le problème d'interpolation SR n'est pas inversible, la super-résolution n'est pas faisable sans régularisation. Comme notre objectif est de minimiser la quantité de régularisation, on commence par étudier les conditions critiques pour l'inversibilité en termes de nombre d'images.

On pourra ensuite étudier la reconstruction SR lorsque le problème est inversible. On donne une condition suffisante sur le nombre d'images et les mouvements (des grilles

d'échantillonnage) pour l'inversibilité dans les hypothèses définies au début du chapitre.

On nomme les grilles suivantes : $\Gamma^{hr} = [1, Ml]^2 \cap \mathbb{Z}^2$ et $\Gamma = [1, Ml]^2 \cap (M\mathbb{Z})^2$. On définit Γ^c comme le complément de Γ dans Γ^{hr} , c'est-à-dire le support des images dans le noyau de S . On donne la condition suivante d'inversibilité de A : la différence de mouvement entre les positions de Γ^c ne doit pas être un nombre entier, coordonnée par coordonnée.

Théorème 3.1.1. *Si $N \geq M^2$, $F = I$ et pour tout $p_i, p_j \in \Gamma^c, 1 \leq k_1 < k_2 \leq N$, l'identité $\|(q_{k_1}^{-1}p_i - q_{k_2}^{-1}p_j) \bmod 1\|_0 = 2$ est vérifiée (où \bmod est le modulo), alors A est injective.*

Démonstration. On montre par récurrence sur N que, sous ces hypothèses, ajouter une image BR diminue la dimension du noyau de A par un facteur L^2 . Le noyau est réduit à $\{0\}$ si sa dimension est 0, ce qui est le cas lorsque $N \geq M^2$. Pour plus de clarté, on a divisé la preuve en lemmes. On présente ici l'argument principal, les lemmes intermédiaires sont démontrés dans la Section 3.1.3. Soit :

$$\begin{aligned} A_n : (\mathbb{C}^{Ml \times Ml}) &\rightarrow (\mathbb{C}^{l \times l}) \\ u &\rightarrow SQ_n u \end{aligned} \quad (3.6)$$

Alors $\ker A = \cap_{k=1, n} \ker A_k$.

On montre par récurrence : pour tout $1 < n \leq M^2$, $\dim \cap_{k=1, n} \ker A_k = (M^2 - n)l^2$.

Pour $n = 1$: soit $p_i \in \Gamma^c$. Soit $v_i = \mathbf{1}_{p_i}$, la fonction indicatrice du pixel p_i . Soit $u_i = Q_1^{-1}v_i$. On a $Sv_i = 0$. En conséquence, $A_1 u_i = 0$ et $u_i \in \ker A_1$. On vient de définir $(M^2 - 1)l^2$ vecteurs u_i indépendants qui génèrent $\ker A_1$: $\text{vect}(u_i)_{i=1, (M^2-1)l^2} = \ker A_1$. La dimension de l'intersection est donc :

$$\dim \ker A_1 = (M^2 - 1)l^2 \quad (3.7)$$

On traite le cas $n = 2$, pour bien illustrer comment les noyaux des A_k interagissent. Pour $n = 2$: de manière similaire, on construit $\text{vect}(u'_i)_{i=1, (M^2-1)l^2} = \ker A_2$. Avec le Lemme 3.1.2 ($\ker A_1 + \ker A_2 = \mathbb{C}^{Ml \times Ml}$), la dimension de l'intersection est :

$$\begin{aligned} \dim(\ker A_1 \cap \ker A_2) &= \dim(\ker A_1) + \dim(\ker A_2) \\ &\quad - \dim(\ker A_1 + \ker A_2) \\ &= (M^2 - 2)l^2. \end{aligned} \quad (3.8)$$

Soit $n > 2$. On suppose que $\dim \cap_{k=1, n} \ker A_k = (M^2 - n)l^2$. Avec le Lemme 3.1.3, on a $(\cap_{k=1, n} \ker A_k) + \ker A_{n+1} = \mathbb{C}^{Ml \times Ml}$. En utilisant les mêmes relations de dimension que pour $n = 2$, on obtient pour $n + 1$:

$$\begin{aligned} \dim \cap_{k=1, n+1} \ker A_k &= \dim \cap_{k=1, n} \ker A_k + \dim \ker A_{n+1} \\ &\quad - \dim(\cap_{k=1, n} \ker A_k + \ker A_{n+1}) \\ &= (M^2 - n)l^2 + (M^2 - 1)l^2 - M^2 l^2 \\ &= (M^2 - n - 1)l^2, \end{aligned} \quad (3.9)$$

ce qui montre que l'hypothèse de récurrence est vraie au rang $n + 1$. En conclusion pour $N = M^2$, la dimension de l'intersection des noyaux des A_k est

$$\dim \cap_{k=1, n+1} \ker A_k = (M^2 - M^2)l^2 = 0 \quad (3.10)$$

et A est injective. Pour $N > M^2$, il suffit de prendre n'importe quel sous-ensemble de M^2 noyaux de A_k pour obtenir le résultat. \square

La principale condition pour obtenir l'injectivité est le nombre d'images qui est le même que pour les translations par Papoulis. Ce théorème élimine quelques cas inversibles. Des exemples de configurations inversibles et non inversibles sont montrés sur la figure 3.3. En

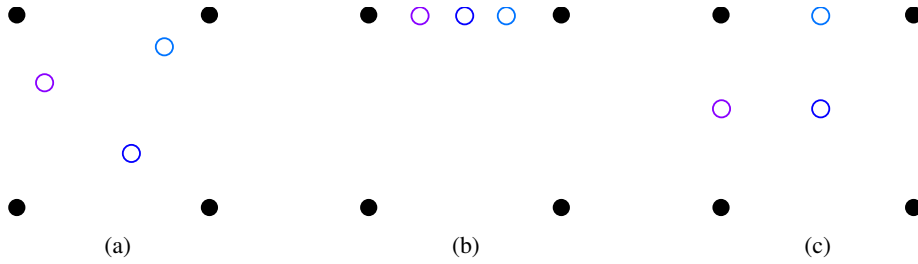


FIGURE 3.3. Configurations inversibles et non inversibles couvertes par Théorème 3.1.1. $M = 2$ et mouvement translationnel. Les points noirs forment la grille de référence (a) Cas inversible dans les hypothèses. (b) Cas non inversible exclu par le théorème. (c) Cas exclu par le théorème, mais inversible.

pratique, c'est la version probabiliste qui nous intéresse. La principale conséquence est que le problème d'interpolation SR avec $N \geq M^2$ est presque sûrement inversible pour des mouvements aléatoires :

Corollaire 3.1.1. *Si $N \geq M^2$ et les mouvements sont aléatoires, A est injective presque sûrement.*

Démonstration. L'espace des paramètres de mouvement $E \subset \mathbb{R}^{6N}$ exclus par le Théorème 3.1.1 a une mesure nulle. En effet, quelque soit $\theta \in \mathbb{R}^6$, l'ensemble $(\mathbb{R} \cdot \theta) \cap E$ est de mesure 0 parce qu'il est dénombrable. En utilisant Fubini, E a mesure 0 dans \mathbb{R}^6 . \square

On peut comparer ce résultat avec [13] où il est montré que le problème d'échantillonnage aléatoire des polynômes trigonométriques est inversible presque sûrement si on dispose d'au moins autant d'équations que d'inconnues (équivalent à $N = M^2$). Notre résultat probabiliste est différent, car les positions d'échantillonnage ne sont pas complètement aléatoires. On a aussi donné une condition déterministe pour cette inversibilité. Par exemple, si le mouvement est une translation dans la seule direction horizontale, l'hypothèse du théorème n'est pas remplie (comme on peut le voir sur la figure 3.3) et ce cas n'est effectivement pas inversible.

3.1.3 Résultats intermédiaires pour l'inversibilité

L'image de la fonction indicatrice d'un pixel par une transformation affine dans le cadre discret peut être vue comme un vecteur de type Vandermonde dans le domaine de Fourier.

Dans ce premier lemme, on étend le résultat de liberté des familles de vecteurs de Vandermonde au cas 2D.

Lemme 3.1.1. *Pour $1 \leq i \leq N$, soit $u_i \in \mathbb{C}^{n \times n}$ tel que $u_i(r, s) = x_i^r y_i^s$. On appelle u_i un vecteur de Vandermonde 2D avec générateur $[x_i, y_i]$. Si pour tout $1 \leq i < j \leq N$, $x_i \neq x_j$, $y_i \neq y_j$, alors $\dim(\text{vect}(u_i)_{i=1,N}) = \min(N, n^2)$.*

Démonstration. On montre que les u_i sont linéairement indépendants si $N \leq n^2$. Supposons que $\sum_{i=1,N} \lambda_i u_i = 0$. Soit $u_i(s) = X_i y_i^s$ avec $X_i = (x_i^r)_r$. Pour tout s , on a donc

$$\sum_{i=1,N} \lambda_i u_i(s) = \sum_{i=1,N} \lambda_i X_i y_i^s = 0. \quad (3.11)$$

Les X_i forment une famille libre de vecteurs de Vandermonde 1D. Ce qui implique que $\sum_{i=1,N} \lambda_i y_i^s = 0$ que l'on peut réécrire $\sum_{i=1,N} \lambda_i Y_i = 0$, avec $Y_i = (y_i^s)_s$. Or les Y_i sont aussi des vecteurs de Vandermonde 1D indépendants. On a donc, pour tout i , $\lambda_i = 0$, et les u_i sont indépendants. \square

Pour pouvoir calculer la dimension de l'intersection des noyaux de A_1 et A_2 , on a besoin de connaître la dimension de la somme de ces espaces. Cette propriété est nécessaire pour l'initialisation de la récurrence du théorème 3.1.1.

Lemme 3.1.2. *Si pour tout $p_i, p_j \in \Gamma^c$, on a l'identité $\|(q_1^{-1} p_i - q_2^{-1} p_j) \bmod 1\|_0 = 2$, alors $\ker A_1 + \ker A_2 = \mathbb{C}^{Ml \times Ml}$.*

Démonstration. On prouve ce lemme pour des mouvements affines Q_i sur des signaux discrets finis. D'après l'hypothèse 4, on peut appliquer Q_i^{-1} et Q_i par interpolation de Fourier discrète finie. On peut construire une base de $\ker A_1$ et $\ker A_2$ en prenant l'inverse des transformations des fonctions indicatrices des pixels mis à zéro par le sous-échantillonnage. Dans le domaine de Fourier, ces bases sont :

$$\hat{u}_i(\omega) = e^{-j\langle \omega, q_1^{-1} p_i \rangle}, \hat{u}'_i(\omega) = e^{-j\langle \omega, q_2^{-1} p_i \rangle}. \quad (3.12)$$

Ce sont des vecteurs de Vandermonde 2D avec générateurs $[e^{-j\langle e_x, q_1^{-1} p_i \rangle}, e^{-j\langle e_y, q_2^{-1} p_i \rangle}]$, avec $e_x = [1, 0]^T$ et $e_y = [0, 1]^T$. Les hypothèses que pour tout p_i, p_j , $\|q_1^{-1} p_i - q_2^{-1} p_j\|_0 \bmod 1$ implique que les générateurs de Vandermonde 2D sont différents 2 à 2. On utilise alors le Lemme 3.1.1 : $\ker A_1 + \ker A_2 = \text{vect}((\hat{u}_i), (\hat{u}'_i)) = \mathbb{C}^{Ml \times Ml}$. \square

On refait le même raisonnement pour l'hypothèse de récurrence du théorème.

Lemme 3.1.3. *Soit $n < M^2$. Si pour tout $p_i, p_j \in \Gamma^c$ et pour tout $1 \leq k_1 < k_2 \leq N$, on a l'identité $\|q_{k_1}^{-1} p_i - q_{k_2}^{-1} p_j \bmod 1\|_0 = 2$ et $\dim(\cap_{k=1,n} \ker A_k) = (M^2 - n)l^2$ alors $\cap_{k=1,n} \ker A_k + \ker A_{n+1} = \mathbb{C}^{Ml \times Ml}$.*

Démonstration. Soit (\hat{b}_i) une base de $\cap_{k=1,n} \ker A_k$ de taille $(M^2 - n)l^2$ dans le domaine de Fourier. Dans la base $(\hat{u}_j)_{j=1,n}$ de $\ker A_1$, :

$$\hat{b}_i = \sum \alpha_{i,j} \hat{u}_j \quad (3.13)$$

Soit \hat{u}'_i une base de $\ker A_{n+1}$ dans le domaine de Fourier. Avec l'hypothèse, toute combinaison linéaire de \hat{b}_i, \hat{u}'_i est une combinaison linéaire de vecteurs de Vandermonde 2D indépendants. Alors, $\dim(\text{vect}((\hat{b}_i), (\hat{u}'_i))) = \min((Ml)^2, (M^2 - n)l^2 + (M^2 - 1)l^2) = (Ml)^2$. Finalement, on a $\cap_{k=1,n} \ker A_k + \ker A_{n+1} = \text{vect}((\hat{b}_i), (\hat{u}'_i)) = \mathbb{C}^{Ml \times Ml}$. \square

3.2 Conditionnement en fonction du nombre d'images

On étend le comportement asymptotique de l'interpolation SR bruitée étudié par Robinson [90] et Champagnat [26] à notre modèle et aux mouvements affines. Ces deux travaux montrent expérimentalement et théoriquement que l'erreur de reconstruction de la SR par minimisation L^2 tend vers 0 lorsque le nombre d'images tend vers l'infini. On sépare cependant dans notre étude l'effet de l'inversion du système et l'effet de débruitage dû à la multiplication du nombre d'images. On montre de plus qu'il est toujours intéressant de prendre plus d'images BR. On étudie ensuite l'estimation des paramètres de mouvement affine pour un grand nombre d'images en calculant la matrice hessienne (ou simplement hessienne) de la méthode d'estimation par moindres carrés non linéaires appelée projection variable. Cela permet de comprendre le comportement asymptotique de cette méthode et montre qu'un terme de régularisation (comme dans [88]) n'est pas toujours nécessaire.

3.2.1 Interpolation

On suppose que $N \geq M^2$. En général, les données capturées sont contaminées par du bruit additif n et on observe les images BR $w = Au_0 + n$. L'image HR estimée par minimisation L^2 est alors contaminée par un bruit de reconstruction. En effet, lorsque l'on minimise pour A fixe et $\lambda = 0$ le problème (2.8) :

$$\tilde{u} = \operatorname{argmin}_u \|Au - w\|_2^2. \quad (3.14)$$

On obtient la solution unique des moindres carrés

$$\tilde{u} = A^\dagger w = (A^H A)^{-1} A^H w. \quad (3.15)$$

$A^\dagger = (A^H A)^{-1} A^H$ est appelée pseudo-inverse de A . Le bruit de reconstruction est donc

$$n_R = A^\dagger n. \quad (3.16)$$

Minimiser ce bruit est alors la principale préoccupation pour que la super-résolution soit bien informée avec la précision voulue. Un bon conditionnement est critique pour que la reconstruction soit de bonne qualité. Le conditionnement de la matrice d'un système linéaire mesure la difficulté de sa résolution et permet de déterminer une borne de l'énergie du bruit en sortie de l'inversion. On définit le conditionnement en norme 2 ainsi :

Définition 3.2.1 (Conditionnement L^2). *On définit le conditionnement L^2 d'une matrice A comme le ratio de ses valeurs singulières extrémales. Soit σ_{max} est la valeur singulière de A la plus grande et σ_{min} la plus petite. Le conditionnement $\kappa(A)$ est alors défini par :*

$$\kappa(A) = \frac{\sigma_{max}}{\sigma_{min}} \quad (3.17)$$

On a de plus la relation $\kappa(A)^2 = \kappa(A^H A)$ (car $N \geq M^2$). Il est utile de connaître le conditionnement, car pour un bruit d'entrée n le bruit de reconstruction n_R a une énergie bornée :

$$\|n_R\|_2^2 \leq \frac{\kappa(A)^4}{\sigma_{max}^4} \|A^H n\|_2^2 \quad (3.18)$$

Plusieurs facteurs influencent le conditionnement de la super-résolution. Baker [11] montre que le conditionnement augmente lorsque le facteur de SR M augmente. Le conditionnement dépend aussi de la distribution de l'échantillonnage [51]. Pour un mouvement contrôlé un conditionnement de 1 est obtenu en espaçant régulièrement les grilles BR afin qu'elles se superposent à la grille HR. Lorsque les mouvements sont aléatoires et uniformes, nous allons montrer que le conditionnement de l'interpolation SR affine tend vers 1 lorsque le nombre d'images augmente, et on en déduit la vitesse de convergence. Cela a été montré expérimentalement pour une SR translationnelle dans [90] en terme de borne de Cramer-Rao pour l'estimation de l'image HR. [26] montre que l'erreur de reconstruction tend vers 0 lorsque le nombre d'images augmente. On suppose que les translations après application de la partie linéaire sont uniformes dans $[0, M]^2$. L'hypothèse sur les translations suppose que l'amplitude des mouvements est assez grande. La taille de chaque capteur est souvent très petite (de l'ordre de 10 microns), ce qui rend la partie fractionnelle du mouvement causé par le mouvement des mains presque aléatoire. De plus, si le mouvement a une variance assez grande et une loi assez lisse, les translations peuvent être considérées comme uniformes.

On illustre ce fait en 1D, sur la figure 3.4. On génère 10 000 échantillons d'une loi gaussienne pour différentes variances, on enlève la partie entière de ces échantillons et génère un histogramme. On remarque que lorsque la variance augmente, la partie fractionnelle se comporte comme une loi uniforme.

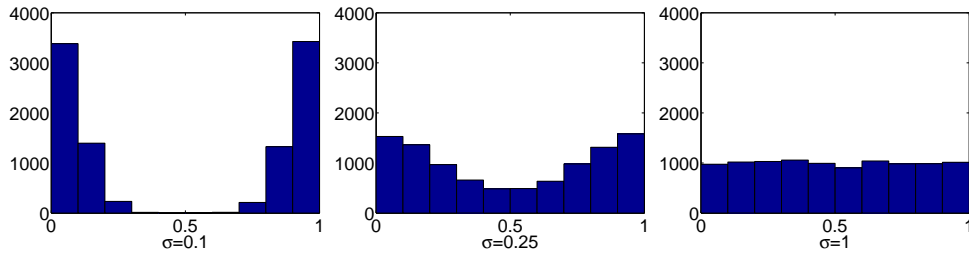


FIGURE 3.4. Validité de l'hypothèse de translation uniforme. Partie fractionnelle d'une variable gaussienne pour différentes variances σ .

Dans ces hypothèses, une image HR peut être parfaitement retrouvée en prenant la pseudo-inverse de A (Corollaire 3.1.1) si le bruit est nul. On montre le théorème suivant en utilisant une technique similaire à [26] :

Proposition 3.2.1. *On suppose que les affinités q_i ont la distribution suivante : les translations $t'_i = l_i^{-1}t_i$ appliquées après la partie linéaire du mouvement affine sont uniformes dans $[0, M]^2$. Alors le conditionnement κ du système tend vers 1 (au sens des distributions) lorsque le nombre d'images augmente.*

Démonstration. Soit $\Gamma = \mathbb{Z}^2$ et $\Gamma' = q_i^{-1}\Gamma$. Pour retrouver u_0 à une pulsation particulière $\omega \in \mathbb{R}^2$, on écrit d'abord la transformée de Fourier de u_0 sur la grille transformée Γ' , $\hat{u}_{\Gamma'}(\omega)$ comme une combinaison linéaire des $\hat{w}_i(\omega)$. On peut ensuite par changement de variable retrouver la transformée de Fourier \hat{u} sur Γ et enfin u_0 . En utilisant les règles de la

transformée de Fourier, on a

$$\begin{aligned}\widehat{w}_i(\omega) &= \widehat{SQ_i u}(\omega) \\ &= \frac{1}{M^2} \sum_{k \in \mathbb{Z}^2} \widehat{Q_i u}(\omega + \frac{2\pi k}{M}).\end{aligned}\quad (3.19)$$

D'autre part, avec

$$\begin{aligned}\widehat{Q_i u}(\omega) &= \sum_{\mathbf{x} \in \Gamma} u(q_i(\mathbf{x})) e^{-j\langle \omega, \mathbf{x} \rangle} \\ &= \sum_{\mathbf{x} \in \Gamma'} e^{-j\langle \omega, q_i^{-1} \mathbf{x} \rangle} u(\mathbf{x}) \\ &= \widehat{u}_{\Gamma'}(l^{-T} \omega) e^{j\langle \omega, l_i^{-1} t_i \rangle}.\end{aligned}\quad (3.20)$$

On peut alors écrire le système d'équations :

$$\widehat{w}_i(l_i^T \omega) = \frac{1}{M^2} \sum_{k \in \mathbb{Z}^2} \widehat{u}(\omega + l_i^{-T} \frac{2\pi k}{M}) e^{j\langle \omega, t_i \rangle + j\langle \frac{2\pi k}{M}, l_i^{-1} t_i \rangle}.\quad (3.21)$$

Seulement M^2 termes sont différents de 0 dans la somme. S'il y a plus de M^2 images BR, c'est un système surdéterminé de taille $N \times M^2$ pour chaque pulsation :

$$C \widehat{u}_{al} = w \quad (3.22)$$

où $\widehat{u}_{al}(k) = \widehat{u}(\omega + l_i^{-T} \frac{2\pi k}{M})$, $w(i) = \widehat{w}_i(l_i^T \omega)$ et $C_{i,k} = \frac{1}{M} e^{j\langle \omega, t_i \rangle + j\langle \frac{2\pi k}{M}, l_i^{-1} t_i \rangle}$. La matrice C est le produit de 2 matrices :

$$C = \Delta B \quad (3.23)$$

avec $\Delta = \frac{1}{M^2} \text{diag}(e^{j\langle \omega, t_i \rangle})$ et $B_{i,k} = e^{j\langle \frac{2\pi k}{M}, l_i^{-1} t_i \rangle}$. Le conditionnement du système est alors le conditionnement de $R = B^H B$ qui est une matrice de Toeplitz de terme général $R_{r,s} = \sum_i e^{j\langle \frac{2\pi(s-r)}{M}, t_i \rangle}$. On peut montrer avec une application directe du théorème centrale limite que R converge en distribution vers un multiple de l'identité, car les nombres $e^{j\langle \frac{2\pi(s-r)}{M}, l_i^{-1} t_i \rangle}$ convergent vers une distribution uniforme sur le cercle unité (pour $s \neq r$) dont la moyenne est nulle. Par continuité du conditionnement, $\kappa(R)$ converge vers 1. \square

Un grand nombre d'images est donc statistiquement meilleur. Le système tend vers un système parfaitement inversible, auquel s'ajoute l'effet de moyennage dû à la quantité d'observations. De l'application du théorème central limite, on peut déduire l'inégalité de concentration suivante sur $R' = \frac{1}{N} R$. La probabilité que les termes non diagonaux ($r \neq s$) soient d'énergie supérieure à λ est :

$$P(|R'(r, s)| \geq \lambda) \leq 2\phi(-\lambda\sqrt{N}) \quad (3.24)$$

où ϕ est la fonction de répartition de la distribution gaussienne. Le théorème de Gershgorin [69] permet d'encadrer les valeurs propres de $R' - I$. On a :

$$\sigma_k(R' - I) \leq (M^2 - 1) \max_{r \neq s} |R'(r, s)| \quad (3.25)$$

On a donc en bornant grossièrement l'intersection des probabilités (pour la répartition du max) par la probabilité d'un évènement (avec un changement de variable sur λ) :

$$P(|\sigma_{max}(R') - 1| \geq \lambda \text{ ou } |\sigma_{min}(R') - 1| \geq \lambda) \leq 2\phi(-c_0\lambda\sqrt{N}) \quad (3.26)$$

ou $c_0 = 1/(M^2 - 1)$ et

$$P(|\kappa(R') - 1| \geq 2\lambda/(1 - \lambda)) \leq 2\phi(-c_0\lambda\sqrt{N}). \quad (3.27)$$

À probabilité fixée, le conditionnement tend vers 1 à une vitesse $O(1/\sqrt{N})$. Grâce à cette vitesse de convergence, on peut déduire, qu'étant donné une erreur, avec une probabilité donnée, il existe des constantes c_1, c_2 telles que l'erreur de reconstruction est bornée par :

$$\|n_R\|_2^2 \leq \underbrace{c_1 \left(1 + c_2 \frac{1}{\sqrt{N}}\right)^4}_{\text{Inversion}} \underbrace{\frac{1}{N^2} \|A^H n\|_2^2}_{\text{Compositage}}. \quad (3.28)$$

Si l'énergie du bruit tend vers 0 lorsqu'on le moyenne (par «compositage», voir ci-dessous le cas du bruit gaussien), alors le bruit de reconstruction tend vers 0 lorsque N tend vers $+\infty$. Il suffit que le bruit n_i qui contamine chaque image BR soit de même énergie, à moyenne nulle et indépendant du bruit des autres images. On a alors :

$$\begin{aligned} E\|A^H n\|_2^2 &= E\left\|\sum_i Q_i S^H n_i\right\|_2^2 \\ &= \sum_i E\|Q_i S^H n_i\|_2^2 \\ &\leq \sum_i E\|n_i\|_2^2 \\ &\leq NE\|n_1\|_2^2 \end{aligned} \quad (3.29)$$

La borne de l'équation (3.28) devient alors

$$E\|n_R\|_2^2 \leq \underbrace{c_1 \left(1 + c_2 \frac{1}{\sqrt{N}}\right)^4}_{\text{Inversion}} \underbrace{\frac{1}{N} E\|n_1\|_2^2}_{\text{Compositage}} \quad (3.30)$$

Dans la figure 3.5, on effectue 10 interpolations SR affines bruitées par un bruit gaussien avec paramètres aléatoires pour différents N . On observe la convergence de l'erreur vers 0 lorsque $N \rightarrow \infty$, la convergence théorique calculée semble être respectée.

Dans certains cas, le bruit de reconstruction par minimisation L^2 ne tend pas vers 0. Si le bruit qui contamine les observations a une composante continue dans une certaine zone de la scène sur plusieurs images de basse résolution. On peut penser par exemple à un oiseau qui s'envole, une voiture qui se gare. La minimisation L^2 n'est pas bien adaptée à ces bruits. On s'intéressera à la manière d'être robuste à ces bruits dans les chapitres 6 et 7.

Dans la figure 3.6, on montre l'évolution du logarithme du conditionnement de la SR dans le cas translationnel 1D (où il est facilement calculable). On observe la convergence vers 0 de ce logarithme. On voit comme dans [11] que ce conditionnement augmente avec M . On a aussi dessiné $\log(\kappa - 1)$ avec la vitesse de convergence théorique de $\log(1/\sqrt{N})$.

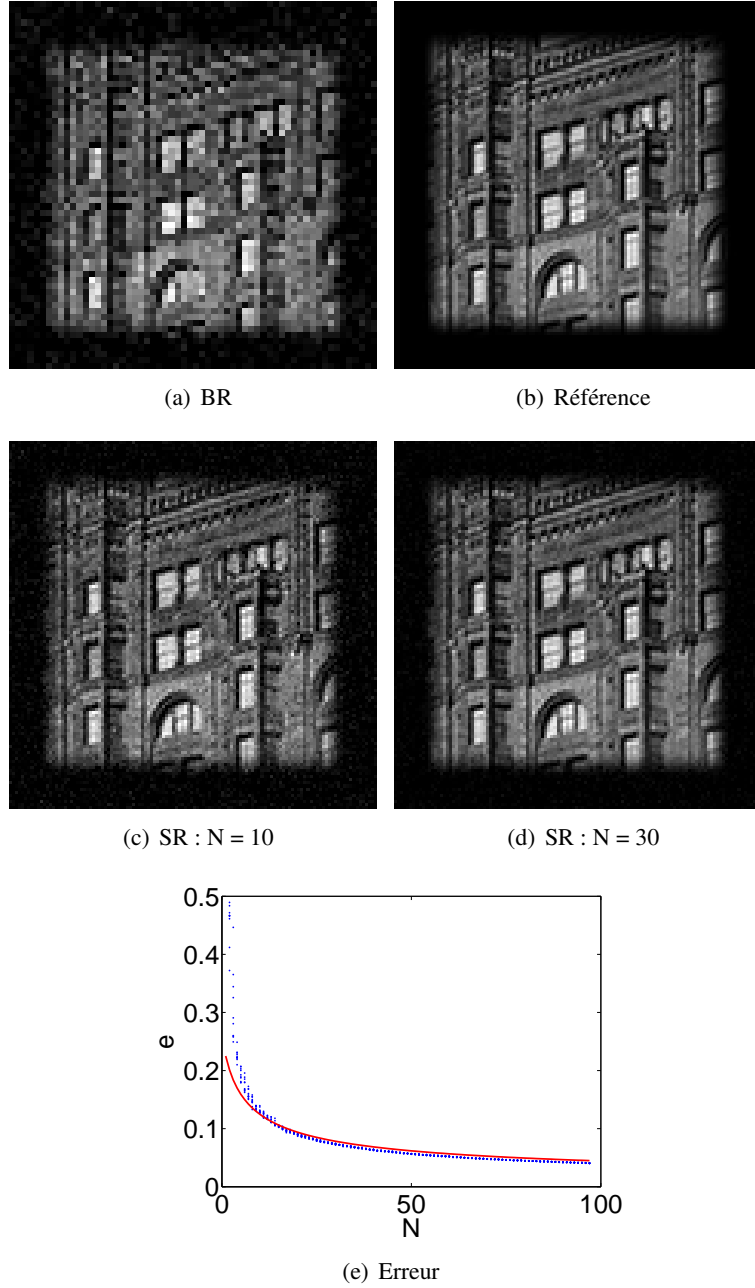


FIGURE 3.5. Erreur de reconstruction de la SR bruitée en fonction de N (a) une image BR, (b) l'image HR (c) résultat de la SR pour $N=10$ (d) résultat de la SR pour $N=30$, (e) erreur de reconstruction en fonction de N , en bleu expériences, en rouge, vitesse de convergence théorique.

Ce résultat permet de déduire une méthode de reconstruction par compositage que l'on peut rapprocher de la méthode «shift and add» [44]. Comme $\kappa(R) \rightarrow 1$, $R \sim N.Id$ et $A^\dagger \sim \frac{1}{N}A^H$. On peut utiliser $\frac{1}{N}A^H$ comme un opérateur de reconstruction lorsque beaucoup d'images sont disponibles. On représente figure 3.7 l'erreur de reconstruction

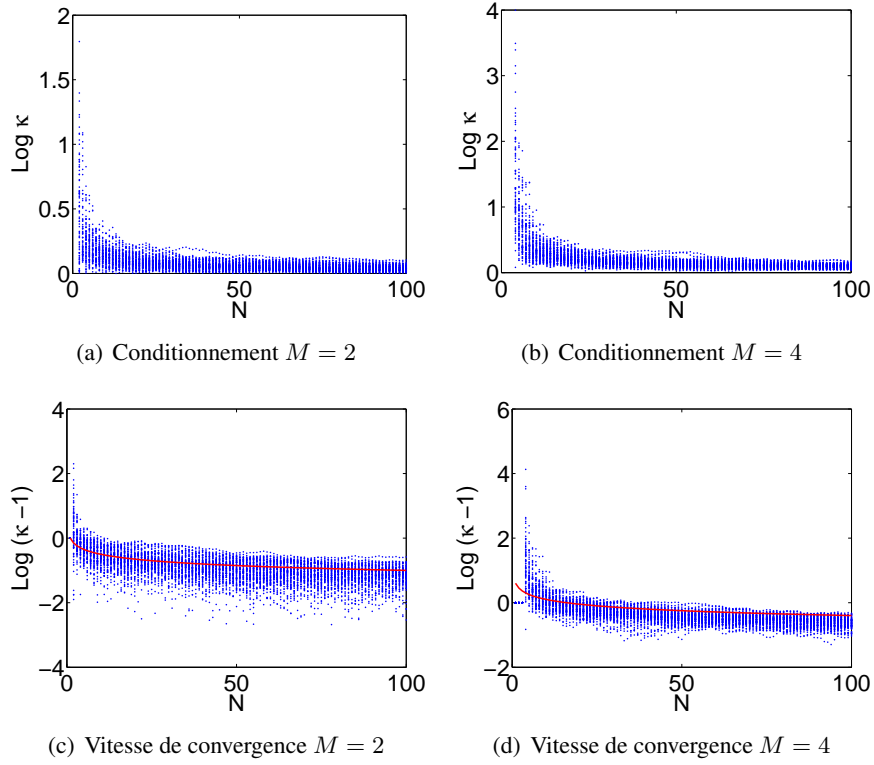


FIGURE 3.6. Convergence du conditionnement pour la SR 1D. (a) Expériences avec $M = 2$. (b) Expériences avec $M = 4$. (c) Vitesse de convergence avec $M = 2$ (vitesse théorique en rouge). (d) Vitesse de convergence avec $M = 4$ (vitesse théorique en rouge).

$e_N = \|u_0 - u_N\|$ de l'estimateur $u_N = \frac{1}{N} A^H w$ par rapport à N . Pour chaque valeur de N , on génère 100 expériences où les mouvements affines ont des translations, des rotations, et des zooms uniformes. La même image HR (que la figure. 3.2) est utilisée pour toutes les expériences. Dans la figure 3.8, on effectue la même expérience pour le cas affine, en ajoutant un bruit gaussien de variance constante aux images.

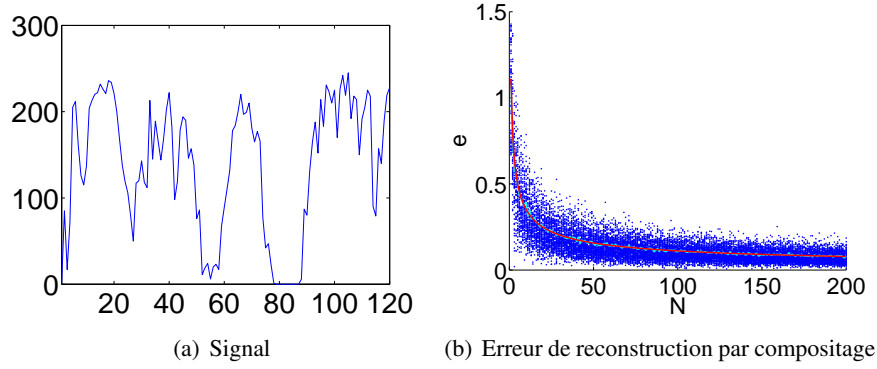


FIGURE 3.7. Convergence de l'estimateur de compositage $\frac{1}{N}A^H w$ dans le cas translationnel 1D. (a) Signal de référence. (b) Erreur de reconstruction (bleu clair) avec $M=3$ avec la vitesse de convergence théorique en rouge. En bleu clair, on montre la moyenne de l'erreur pour chaque valeur de N (courbe cachée sous la prédiction).

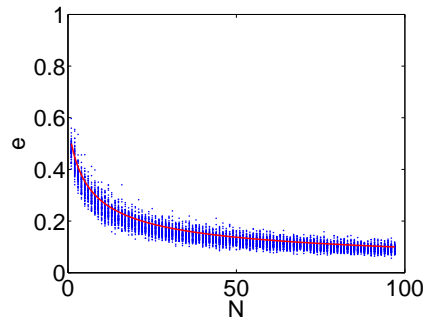


FIGURE 3.8. Convergence de l'estimateur de compositage $\frac{1}{N}A^H w$ dans le cas affine bruité. Erreur de reconstruction par rapport au nombre d'images BR.

On montre maintenant dans un contexte déterministe qu'il est toujours intéressant d'ajouter une observation (une image BR). Autrement dit, le bruit de reconstruction décroît avec le nombre d'images. Ici $n \sim N(0, \sigma^2 I)$ est un bruit blanc gaussien de moyenne nulle. On a donc $n_R \sim N(0, \sigma^2 A^\dagger A^{\dagger H}) = N(0, \sigma^2 (A^H A)^{-1})$. Ce bruit de reconstruction aura une énergie normalisée $e = \sigma^2 \text{tr}((A^H A)^{-1}) = \sum_i \frac{1}{\lambda_i}$ où les λ_i sont les valeurs propres (v.p.) de $A^H A$.

On ajoute une image :

$$\begin{aligned} A' : (\mathbb{C}^{Ml \times Ml}) &\rightarrow (\mathbb{C}^{l \times l})^{N+1} \\ u &\rightarrow (SQ_k u)_{k=1, N+1} \end{aligned} \quad (3.31)$$

$$w' = A' u_0 + n' \quad (3.32)$$

avec $n' \sim N(0, \sigma'^2 I)$. On a : $e' = \sigma'^2 \text{tr}((A'^H A')^{-1}) = \sum \frac{1}{\sigma'^2}$ avec $\sigma'^2 = v.p.(A'^H A')$.

Proposition 3.2.2. *Capter plus d'images diminue le bruit, autrement dit $e' \leq e$*

Démonstration. On montre d'abord que $\lambda'_i \geq \lambda_i$ pour tout i . En utilisant les inégalités de Weyl :

$$\lambda_i \leq \lambda'_i + \lambda_{\max}(A^H A - A'^H A') \quad (3.33)$$

On a :

$$\begin{aligned} A'^H A' - A^H A &= \sum_{k=1, N+1} Q_k^H S^H S Q_k - \sum_{k=1, N} Q_k^H S^H S Q_k \\ &= Q_{N+1}^H S^H S Q_{N+1} \end{aligned} \quad (3.34)$$

qui est une application linéaire positive symétrique (c'est-à-dire que toutes ses valeurs propres sont positives). Alors $\lambda_{\max}(A^H A - A'^H A') \leq 0$ et $\lambda_i \leq \lambda'_i$. Avec ce résultat, on a $\frac{1}{\lambda'_i} \leq \frac{1}{\lambda_i}$ pour tout i . En conséquence :

$$\begin{aligned} \sum_i \frac{1}{\sigma'^2} &\leq \sum_i \frac{1}{\lambda_i} \\ \sigma'^2 \sum_i \frac{1}{\lambda'_i} &\leq \sigma^2 \sum_i \frac{1}{\lambda_i} \\ e' &\leq e. \end{aligned} \quad (3.35)$$

□

Ce résultat montre que si l'on est capable de retrouver les paramètres d'acquisition, la meilleure stratégie est de garder toutes les images BR disponibles pour effectuer la reconstruction. Ce résultat est en accord avec l'intuition suivante : avoir plus de données augmente le rapport signal sur bruit. Il faut aussi noter que cette propriété dépend de la structure de l'application linéaire qui génère les données.

3.2.2 Estimation des paramètres de mouvement

S'il peut sembler naturel qu'augmenter le nombre d'observations d'un système linéaire bruité diminue le bruit dans le cas de l'interpolation SR, il est moins évident de savoir si l'estimation des mouvements se comporte aussi bien. On s'attache à répondre à cette question dans cette section.

Projection variable et convergence

Lorsque l'on effectue la reconstruction SR, il faut estimer les mouvements pour chaque image BR. Dans le cas d'un mouvement affine, il faut estimer 6 paramètres par image BR. Des techniques d'estimation entre deux images bien échantillonnées existent [86]. On peut citer les méthodes à base de descripteurs [16, 97] ou de flot optique [46].

Cependant, dans le cas de la SR, les images sont repliées (aliasing), et ces méthodes ne donnent pas une estimation assez précise pour effectuer une bonne reconstruction [117, 89, 90]. Elle peuvent cependant être utiles pour déterminer une solution initiale.

Dans le cas des translations, Vandewalle *et coll.* [109] proposent une méthode basée sur l'analyse de sous-espaces engendrés par l'image reconstruite pour recalculer les images. Son propos est le suivant : si les images sont recalées, alors elles ne peuvent pas générer un sous-espace plus grand que celui de l'image HR (dont la dimension est supposée connue par ailleurs). Cependant, cette méthode est décrite comme ayant un temps d'exécution long. Sbaiz *et coll.* [93] proposent dans le cas des signaux 1D, une méthode exacte, mais de coût prohibitif, basée sur la réécriture du problème comme une équation polynomiale et l'utilisation de bases de Groebner.

Si le problème d'interpolation est inversible, la reconstruction SR peut être vue comme un problème de minimisation non linéaire en fonction des paramètres de mouvement. Cette méthode s'appelle la projection variable [48]. [88] montre qu'utiliser une version régularisée de la méthode de projection variable donne des résultats stables en pratique. Dans [92], la convergence d'une version de cet algorithme, approchée avec des B-Splines, est aussi observée en pratique. Cependant, la nécessité d'une telle régularisation n'est pas mise en cause. On cherche donc à savoir s'il est possible d'estimer les paramètres sans régularisation lorsque le nombre d'images augmente.

Dans le cas non bruité $w = Au_0$, lorsque $N > M^2$, $u_0 = (A^H A)^{-1} A^H w = A^\dagger(\theta_0)w$. Les vecteurs $\theta_0 \in \mathbb{R}^{6N}$ sont les paramètres des affinités (que l'on décompose comme une translation suivie d'une séquence rotation-zoom 2D anisotrope-rotation (SVD de la partie linéaire)). On utilise la méthode de projection variable pour estimer θ_0 [48, 91]. On minimise :

$$G(\tilde{\theta}) = \|A(\tilde{\theta})A^\dagger(\tilde{\theta})w - w\|_2^2, \quad (3.36)$$

ce qui n'est pas un problème convexe. On a bien $G(\theta_0) = 0$. Si l'image présente des invariances par translation ou une périodicité, ce minimum global ne sera pas unique. Avec une première estimation (qui peut être obtenue avec une méthode de recalage dédiée et que l'on supposera correcte, c'est à dire proche du minimum global), on peut minimiser cette fonctionnelle à l'aide d'une descente de gradient. La vitesse de convergence et la précision de cette méthode dépendent directement du conditionnement de la hessienne H (matrice des dérivées secondes partielles) au point θ_0 (on suppose que la première image n'est pas translatée, et que H n'est pas singulière).

La vitesse de convergence d'une méthode de descente de gradient est bonne lorsque la sensibilité de la fonction (au minimum) à ses différentes variables est semblable (pas de vallée). La hessienne donne l'approximation quadratique d'une fonction à son minimum (car le gradient est nul en ce point). Sa diagonalisation donne un repère orthonormal de l'espace, où chaque direction a une courbure correspondant au coefficient diagonal associé. Si le conditionnement de la hessienne tend vers 1, ses coefficients deviennent égaux et les courbures également. Soit θ_k l'estimation des paramètres à l'étape k de la descente de gradient avec pas optimal (qui minimise la fonctionnelle G). Alors dans le cas où la fonction est parfaitement quadratique on a (voir [63], Chapitre 8) :

$$G(\theta_{k+1}) - G(\theta_0) \leq \left(\frac{\kappa(H) - 1}{\kappa(H) + 1} \right)^2 G(\theta_k) - G(\theta_0) \quad (3.37)$$

La sensibilité au bruit dépend aussi de ces courbures. Comme au minimum, la fonction G se comporte comme une forme quadratique, on obtient que le déplacement de ce minimum est directement dépendant du conditionnement de la hessienne (comme dans la section 3.2.1).

Conditionnement de la matrice hessienne

On calcule conditionnement de la hessienne. Le gradient a l'expression suivante :

$$\begin{aligned} \frac{\partial}{\partial \theta_i} G(\theta) &= 2(A(\theta)A^\dagger(\theta)w - w)^H \left(\frac{\partial}{\partial \theta_i} [A(\theta)A^\dagger(\theta)]w \right) \\ &= 2(A(\theta)A^\dagger(\theta)w - w)^H \left(\frac{\partial}{\partial \theta_i} [A(\theta)]A^\dagger(\theta)w \right. \\ &\quad \left. + A(\theta) \frac{\partial}{\partial \theta_i} [A^\dagger(\theta)]w \right) \\ &= 2(A(\theta)A^\dagger(\theta)w - w)^H \left(\frac{\partial}{\partial \theta_i} [A(\theta)]A^\dagger(\theta) \right) \end{aligned} \quad (3.38)$$

où la dernière ligne est obtenue par orthogonalité. La hessienne est alors (on ne calcule pas la constante) :

$$\begin{aligned} H_{i,j} &\propto \left(\frac{\partial}{\partial \theta_j} [A(\theta)A^\dagger(\theta)]w \right)^H \left(\frac{\partial}{\partial \theta_i} [A(\theta)A^\dagger(\theta)]w \right) \\ &\quad + (A(\theta)A^\dagger(\theta)w - w)^H \left(\frac{\partial^2}{\partial \theta_j \partial \theta_i} [A(\theta)]A^\dagger(\theta)w \right) \end{aligned} \quad (3.39)$$

En particulier, la hessienne évaluée à θ_0 est :

$$H_{i,j} \propto \left(\frac{\partial}{\partial \theta_j} [A(\theta_0)A^\dagger(\theta_0)]w \right)^H \left(\frac{\partial}{\partial \theta_i} [A(\theta_0)]A^\dagger(\theta_0)w \right) \quad (3.40)$$

On prend $\theta = \theta_0$ (on écrit $A = A(\theta_0)$), comme notre objectif principal est de calculer le conditionnement de la hessienne au minimum de la fonctionnelle. Les calculs développés dans la section 3.2.4 donnent l'expression suivante :

$$H_{i,j} \propto \left\langle \frac{\partial}{\partial \theta_j} [A]u_0 \middle| (I - AA^\dagger) \frac{\partial}{\partial \theta_i} [A]u_0 \right\rangle \quad (3.41)$$

On pose $\frac{\partial}{\partial \theta_j}[A]u_0 = w'_j$:

$$H_{i,j} \propto \langle w'_j, (I - AA^\dagger)w'_i \rangle \quad (3.42)$$

Parce que les images BR sont séparées :

$$H_{i,j} \propto \gamma_{i,j} \langle w'_j, w'_i \rangle - \langle A^H w'_j, A^\dagger w'_i \rangle \quad (3.43)$$

où $\gamma_{i,j} = 1$ lorsque θ_i et θ_j sont des paramètres reliés aux mouvements de la même image, et $\gamma_{i,j} = 0$ autrement. Les coefficients $\gamma_{i,j} < w'_j, w'_i \rangle$ définissent une matrice H_B symétrique diagonale par blocs. De plus, les blocs sont de taille 6×6 et bornés au sens de la norme 2. D'après la section précédente, on a $A^\dagger \sim \frac{1}{N}A^H$. Ce qui implique avec l'inégalité de Cauchy Schwartz :

$$\begin{aligned} |\langle A^H w'_j, A^\dagger w'_i \rangle| &= O\left(\frac{1}{N}\right) |\langle A^H w'_j, A^H w'_i \rangle| \\ &= O\left(\frac{1}{N}\right) \|A^H w'_j\| \|A^H w'_i\|. \end{aligned} \quad (3.44)$$

Les quantités $\|A^H w'_j\|$ sont bornées, car $A^H w'_j = Q_j^H S^H w'_j$. On a donc :

$$|\langle A^H w'_j, A^\dagger w'_i \rangle| = O\left(\frac{1}{N}\right) \quad (3.45)$$

et

$$H_{i,j} \propto \gamma_{i,j} \langle w'_j, w'_i \rangle (1 - O\left(\frac{1}{N}\right)) \sim_{N \rightarrow \infty} \gamma_{i,j} \langle w'_j, w'_i \rangle. \quad (3.46)$$

On obtient finalement : $H \rightarrow H_B$ et $\kappa(H) \rightarrow \kappa(H_B)$ par continuité du conditionnement. Calculons l'un des blocs 6×6 de H_B qui correspond à une image BR k_0 . On suppose que θ_j et θ_i sont 2 paramètres du mouvement $Q = Q_{k_0}$. On a :

$$\begin{aligned} H_{Bi,j} &\propto \left\langle \frac{\partial}{\partial \theta_j}[A]u_0 \left| \frac{\partial}{\partial \theta_i}[A]u_0 \right. \right\rangle \\ &\propto \left\langle S \frac{\partial}{\partial \theta_j}[Q]u_0 \left| S \frac{\partial}{\partial \theta_i}[Q]u_0 \right. \right\rangle. \end{aligned} \quad (3.47)$$

H_B représente le conditionnement relatif à l'estimation des paramètres d'une affinité. On a calculé expérimentalement les valeurs propres de 100 H_B différentes en utilisant l'équation (3.47). L'image u_0 a été générée comme un processus gaussien i.i.d. pour avoir un signal qui comporte toutes les fréquences, qui est décorrélié des paramètres d'acquisition. Cela mène à un conditionnement pour H_B plus faible que 100. Comparé à une SR 1D translationnelle où H_B est diagonale avec un conditionnement de 1, le cas plus général de l'affinité est plus difficile à cause de la difficulté intrinsèque de l'estimation de mouvement. En pratique, cette difficulté est limitée, comme on le montre dans l'expérience qui suit.

Pour produire la Table 3.1, on a calculé la moyenne de l'erreur de reconstruction de 10 expériences avec estimation de paramètres en fonction du nombre d'images. On a utilisé un algorithme de gradient conjugué non linéaire, arrêté après le même nombre d'itérations (5 itérations). Pour chaque expérience, le point de départ $\hat{\theta}_d$ de l'algorithme est une perturbation aléatoire de la solution (les translations varient aléatoirement avec un écart type de 1/5ème de pixel, de 3 degrés pour les rotations, et 1 pourcent pour le zoom).

N	4	12	40
$ \tilde{u}_d - u_0 / u_0 $	3.19	0.2	0.1
$ \tilde{u} - u_0 / u_0 $	2.5	0.06	0.04

TABLE 3.1. Erreur de reconstruction moyenne en fonction du nombre d'images. $\tilde{u}_d = A(\tilde{\theta}_d)^\dagger w$

3.2.3 Conclusion sur le conditionnement et expériences réelles

Dans cette section, on a étudié l'évolution du conditionnement en fonction de N . Pour la SR, l'interpolation et l'estimation des paramètres jouent chacune un rôle dans la difficulté du problème. Lorsque l'on fait une hypothèse de mouvement aléatoire avec une distribution réaliste, la partie interpolation tend vers un système parfaitement inversible. L'estimation du mouvement est plus difficile car l'estimation de chaque affinité a un conditionnement intrinsèque. Ce conditionnement dépend du contenu fréquentiel de l'image et de la valeur des paramètres de mouvements. Si le mouvement est translationnel, le conditionnement est de 1 pour des signaux pleine bande. Dans le cas général, ce conditionnement devient bon expérimentalement lorsque les mouvements ne sont pas dégénérés. Il serait intéressant d'étudier plus en détail le problème du recalage des images. Maintenant que l'on sait que l'on peut bien estimer les mouvements à partir d'une solution initiale, il faudrait étudier la taille des minima locaux (c'est à dire l'ensemble de paramètre pour lesquels le minimum est valable) pour ensuite déterminer la finesse de l'estimation initiale dont on a besoin. Une telle analyse a été effectuée dans le cadre du recalage multi-échelle d'images [111].

On montre figure. 3.9 le résultat de deux expériences de SR sans régularisation. On a utilisé des images qui n'ont pas de repliement spectral dans la grille native pour pouvoir simuler une image de vérité terrain. La première expérience est une série de photographies d'une image de ville imprimée. La deuxième est une série de photographies de feuilles en extérieur. Pour chaque expérience, on a pris 20 images avec un appareil photo tenu à la main. On a sous-échantillonné ces images par un facteur 2 et on a effectué la reconstruction sans régularisation avec $M = 2$. On compare le résultat avec l'image de référence qui est une version bruitée de la vérité terrain. On observe que tous les détails ont été reconstruits parfaitement, particulièrement sur les bords repliés, et que le bruit a été réduit.

On voit donc qu'avec une rafale de photographies qui acquiert un nombre d'images réalisable, avec des scènes immobiles, on peut effectuer la super-résolution sur des données réelles sans régularisation.



FIGURE 3.9. Deux expériences de super-résolution sans régularisation avec données réelles. (a) (b) image BR pour les expériences 1 et 2, (c) (d) images de référence (vérité terrain + bruit) (e) (f) images reconstruites sans régularisation.

3.2.4 Calculs intermédiaires pour la hessienne

En utilisant les règles de dérivations usuelles, on a au point θ_0 :

$$\frac{\partial}{\partial \theta_i} [AA^\dagger] = \frac{\partial}{\partial \theta_i} [A]A^\dagger + A \frac{\partial}{\partial \theta_i} [A^\dagger] \quad (3.48)$$

$$= \frac{\partial}{\partial \theta_i} [A]A^\dagger - A(A^H A)^{-1} \frac{\partial}{\partial \theta_i} [A^H A]A^\dagger + A(A^H A)^{-1} \frac{\partial}{\partial \theta_i} [A^H] \quad (3.49)$$

$$= \frac{\partial}{\partial \theta_i} [A]A^\dagger - AA^\dagger \frac{\partial}{\partial \theta_i} [A]A^\dagger - (A^\dagger)^H \frac{\partial}{\partial \theta_i} [A^H]AA^\dagger + (A^\dagger)^H \frac{\partial}{\partial \theta_i} [A^H] \quad (3.50)$$

$$= (I - AA^\dagger) \frac{\partial}{\partial \theta_i} [A]A^\dagger + (A^\dagger)^H \frac{\partial}{\partial \theta_i} [A^H] (I - AA^\dagger) \quad (3.51)$$

$$= 2\text{Re}((I - AA^\dagger) \frac{\partial}{\partial \theta_i} [A]A^\dagger) \quad (3.52)$$

CHAPITRE 4

APPLICATION À LA SUPER-RÉSOLUTION À HAUTE DYNAMIQUE

Le travail décrit dans cette partie a été réalisé en collaboration avec Cecilia Aguerreberre (Telecom Paristech). On y utilise les résultats obtenus sur le conditionnement de la reconstruction SR pour déterminer une quantité de données suffisante pour effectuer une super-résolution et une imagerie haute dynamique simultanément.

La plage dynamique que le capteur d'un appareil photo peut capturer est limitée. Les scènes qui présentent une forte dynamique donneront des images saturées ou sous-exposées. Une façon d'éviter ces effets est de prendre plusieurs photographies de la scène avec des temps d'exposition différents. On appelle ce processus imagerie haute dynamique (High Dynamic Range (HDR) imaging). Dans cette section, on effectue simultanément cette imagerie avec la super-résolution.

Plusieurs algorithmes de génération d'images HDR ont été proposés depuis les travaux de Mann et Picard [67]. Le principe repose sur le calcul d'une moyenne pondérée, prenant en compte les temps d'exposition des différentes acquisitions de chaque pixel. Ce que l'on cherche à estimer est l'irradiance réelle de la scène, c'est-à-dire l'intensité de la lumière qui arrive sur le capteur de l'appareil. Plusieurs schémas de pondération ont été proposés [34, 72, 87, 60, 50, 56, 1]. Le point principal à noter est qu'il est impossible d'aligner des images correctement si elles sont affectées par du repliement spectral (on ne peut pas appliquer la transformation géométrique différemment sur les parties repliées ou non). Cela justifie l'utilisation d'une méthode jointe pour produire des images super-résolues de grande dynamique.

Il existe plusieurs travaux sur l'imagerie de super-résolution à haute dynamique (HDR-SR) jointe. La plus conventionnelle est d'utiliser une minimisation des moindres carrés, comme pour la SR, mais en considérant le domaine de l'irradiance [52, 28] ou d'irradiance transformée [14]. D'autres approches [118] alignent les images BR en utilisant un flot optique puis minimisent une énergie régularisée pour obtenir le résultat. Rad *et coll.* [5] ré-alignent les images BR puis utilisent une triangulation de Delaunay et une interpolation bicubique. Le problème a aussi été étudié dans le cadre de la conception de capteurs adaptés [75, 74]. Toutes ces méthodes utilisent une régularisation. Comme précédemment, on se pose la question du nombre d'images pour obtenir une HDR-SR bien informée. On a aussi intégré les derniers résultats de l'imagerie HDR pour déterminer l'algorithme à utiliser. Les résultats théoriques obtenus sont mis en pratique lors d'expériences réelles.

On considère le modèle d'acquisition de l'irradiance suivant :

$$\begin{aligned} A : \quad \mathbb{R}^{Ml \times Ml} &\rightarrow \mathbb{R}^{N \times (l \times l)} \\ u &\rightarrow (w_i)_{i=1, \dots, N} = (\Omega_i G_i S Q_i u)_{i=1, \dots, N}, \end{aligned} \tag{4.1}$$

où $u \in \mathbb{R}^{Ml \times Ml}$ est maintenant une irradiance HR, $G_i \in \mathbb{R}^{(l \times l) \times (l \times l)}$, $G_i = g t_i$ est le gain d'acquisition (que l'on prendra à $g = 1$ sans perte de généralité et pour simplifier les

expressions) et t_i est le temps d'exposition de chaque image BR. $\Omega_i \in \mathbb{R}^{(l \times l) \times (l \times l)}$ est la matrice diagonale qui prend la valeur 1 si le pixel j n'est ni saturé ni sous-exposé pour le temps d'exposition t_i et 0 autrement. On suppose que $\Omega = (\Omega_i)_{i=1, \dots, N}$ est de rang plein (tous les pixels sont illuminés au moins une fois). On fait quelques simplifications pour faciliter l'étude. On suppose que la réponse des capteurs est linéaire. La PSF est négligée comme avant. L'erreur de quantification n'est pas considérée, car elle est négligeable devant les autres bruits (à part dans des conditions de très basse lumière). On considère toujours un modèle de bruit additif n :

$$w = Au_0 + n. \quad (4.2)$$

Un modèle de Poisson avec approximation gaussienne [2] donne un bruit gaussien spatialement variable

$$n = n_u + n_R, \quad (4.3)$$

où n_u est un bruit gaussien de matrice de covariance Σ_u proportionnelle à l'irradiance ($\Sigma_u = \text{diag}(Au_0)$) et n_R est un bruit de lecture gaussien constant de variance σ_R^2 . La matrice de covariance de n est donc $\Sigma = \text{diag}(Au_0) + \sigma_R^2 I$.

4.1 Méthode de reconstruction

On minimise l'attache aux données L^2 :

$$\tilde{u} = \underset{u}{\operatorname{argmin}} \|W^{1/2}(Au - w)\|_2^2. \quad (4.4)$$

La multiplication par $W^{1/2} = \Sigma^{-1/2}$ normalise le bruit à une variance constante, la solution de (4.4) donne l'estimateur non biaisé de variance minimale. Comme les poids W dépendent de l'irradiance u_0 , une procédure itérative est nécessaire. En pratique, une seule itération en utilisant une version filtrée des images BR est suffisante. La méthode itérative de calcul de l'irradiance [50] est presque optimale dans le cadre de la HDR [1]. Si $\Omega_i = I$, $G_i = I$ et $W = I$ (pas de HDR), l'équation (4.4) est la super-résolution L^2 conventionnelle.

4.2 Stratégie pour le choix des paramètres d'acquisition

Pour effectuer pratiquement la HDR-SR, on doit définir le nombre d'images à acquérir ainsi que les temps d'expositions. Pour des temps t_1, \dots, t_N , l'erreur de reconstruction peut être arbitrairement large. Il n'est pas possible de trouver a priori les temps d'exposition qui minimisent l'erreur de reconstruction du problème joint. Notre but est de chercher à savoir s'il est possible de trouver des paramètres qui garantissent la qualité du résultat. Pour cela, on propose de définir une stratégie de choix des paramètres qui combine les résultats d'optimalité des problèmes de HDR et SR séparés.

Stratégie suivie : Pour le problème de HDR, une image HR est obtenue à partir de p images avec des temps t_1, \dots, t_p , ces temps sont choisis pour minimiser l'erreur de reconstruction HDR (procédure détaillée dans [104]). On propose d'utiliser les conditions de SR qui estiment le mieux chacune de ces p images. Cela garantit que la méthode jointe de reconstruction donne un bon résultat.

Étant donné un ensemble de temps d'expositions t_1, \dots, t_p qui minimisent une erreur de reconstruction HDR, on doit considérer 2 cas. Premièrement, s'il n'y a aucune limite sur le temps d'exposition total et le nombre d'images, pour chaque temps $(t_i)_{i=1, \dots, p}$, le mieux est de prendre autant d'images que possible. D'après la section 3.2, l'erreur diminuera en fonction. Une seconde possibilité est que des limites sur les paramètres existent. Dans la partie suivante, on montre que si le temps d'exposition total est limité, il existe un nombre optimal d'images pour minimiser l'erreur de reconstruction.

4.3 Borne d'erreur de reconstruction pour la HDR-SR

Dans cette section, on borne l'erreur de reconstruction de la HDR-SR pour un temps total d'exposition fixé T et un nombre d'images N acquises avec un temps d'exposition égal T/N .

4.3.1 Nombre optimal d'images pour un temps d'exposition total fixe

On étudie le cas sans saturation. On a alors $\Omega_i = I$. Un nombre d'images N avec un temps d'exposition total T sont acquises. Comme le gain est linéaire le problème est équivalent à l'acquisition de N irradiances BR v avec $A' = G^{-1}A$, et un bruit n' de matrice de covariance. Avec $t = T/N$, on a :

$$\Sigma' = \text{diag}(A'u_0/t + \sigma_R^2/t^2). \quad (4.5)$$

Alors,

$$v = A'u_0 + n'. \quad (4.6)$$

A' est alors l'opérateur de super-résolution conventionnel. On peut donc utiliser les résultats des sections précédentes directement. Le problème devient alors :

$$\tilde{u} = \text{argmin}_v \|W'^{1/2}(A'u_0 - w)\|_2^2, \quad (4.7)$$

avec $W' = \Sigma'^{-1}$.

Borne de l'erreur de reconstruction : On montre que l'énergie du bruit de reconstruction n_{rec} est bornée par (calculs développés dans la section 4.5)

$$\|n_{\text{rec}}\|_2^2 \leq \frac{\kappa(N)^2}{\sigma_{\max}^2(A')} \frac{m^2 l^2}{\tilde{m}} (1 + rN) N^2, \quad (4.8)$$

où $\kappa(N)$ est le conditionnement de $A'^H A'$ (qui est une variable aléatoire qui dépend de N selon la section précédente), $m = \sup(u_0) \simeq \sup(A'u_0)$, $\tilde{m} = \inf(u_0) \simeq \inf(A'u_0)$, $r = \sigma_R^2/(Tm)$. Le nombre optimal d'images pour un temps total fixe est alors celui qui minimise $\|n_{\text{rec}}\|_2^2$.

On a montré dans la section précédente que $\kappa(N)$ peut être borné en probabilité par une fonction de N décroissante. De plus, on peut montrer (6.2) que $N/M^2 \leq \sigma_{\max}^2(A') \leq N$. On peut donc, pour minimiser la borne d'erreur, minimiser la fonction

$$f(N) = \kappa(N)^2 (1 + rN) N, \quad (4.9)$$

en fonction de N . $\kappa(N)$ décroît vers 1. En conséquence, f a un minimum global. Le facteur r est proportionnel à σ_R^2 et ne peut être négligé en basse lumière.

4.3.2 Étude expérimentale de $f(N)$

Données synthétiques : Pour illustrer le comportement de $f(N)$, on calcule la courbe pour le cas translationnel 2D. Pour chaque N , on génère 100 paramètres de translation aléatoire et calcule $f(N)$ en calculant explicitement $\kappa(N)$. On fait de même pour la super-résolution affine. Seulement, $\kappa(N)$ est approximé par le rapport $\|n_{\text{rec}}\|^2/\|n_{\text{in}}\|^2$. Dans la figure 4.1, on montre les résultats obtenus avec $M = 2$ et $r = 0.03$ (avec les valeurs réalistes : $\sigma_R^2 = 30$, $T = 1/10$, $m = 10^4$).

On vérifie que le minimum de f n'est pas atteint au cas critique ($N = M^2 = 4$). Il vaut donc mieux prendre plus de M^2 images pour un temps total donné. On observe une dégradation de la performance pour des grands N . Ceci s'explique que chaque image est illuminée par un temps d'exposition plus court. Le bruit de variance σ_R^2 devient prépondérant. La quantité de bruit dans les images BR devient trop grande pour être compensée par l'effet moyenneur de la super-résolution.

Le facteur r peut être vu comme l'inverse de la gamme dynamique de l'irradiance. Lorsque r diminue, le minimum de $f(N)$ est atteint pour un N plus grand. Enfin, on remarque que la borne d'erreur augmente lentement après le minimum. Si on choisit un nombre d'images un peu plus grand que pour le minimum, le résultat sera peu changé.

Données réelles : On vérifie ces résultats sur données réelles. On prend des photos d'une surface plane pour la figure 4.2. On compare l'erreur de reconstruction HDR-SR ($M = 2$) obtenue à temps total constant T , mais pour 2 temps d'expositions, $t_1 = 1/41$ et $t_2 = 1/64$, et 2 différents nombres d'images $N_1 = T/t_1$ et $N_2 = T/t_2$. De la même façon, on compare à nombre d'images fixé. L'erreur de reconstruction montrée figure 4.2 est estimée à partir des parties grises de l'image. Dans cet exemple, pour un nombre fixe d'images, il est plus intéressant d'utiliser le temps d'exposition le plus long 1/41 secondes. Pour un temps total fixé, il est plus intéressant de combiner plus d'images à 1/64 secondes. La différence est visible sur les images .

4.3.3 Discussion et conséquences pratiques

- **Avec un temps d'exposition total fixé**, le nombre d'images N est supérieur au nombre critique (M^2). N ne doit pas être trop grand, car l'erreur augmente à partir d'une certaine valeur.
- **Avec un nombre d'images fixé**, prendre le plus d'images possible est toujours plus intéressant. Cependant, si on se trouve à la valeur optimale de N , on sait que l'on ne pourra pas faire mieux à temps total fixé. C'est cette remarque qui va guider notre choix de N pour les expériences pratiques.

4.4 Expériences réelles

On effectue des expériences comme suit : on calcule des temps d'expositions [104] optimaux au sens de la HDR, à partir d'un histogramme de la scène. Pour chaque temps d'exposition, on prend un nombre d'images proche du minimum de $f(N)$. En pratique pour $M = 2$, le minimum est atteint autour de $N = 20$. De plus, $N = 20$ est un nombre de photos qu'il est facile de prendre pratiquement, ce qui fixe notre choix.

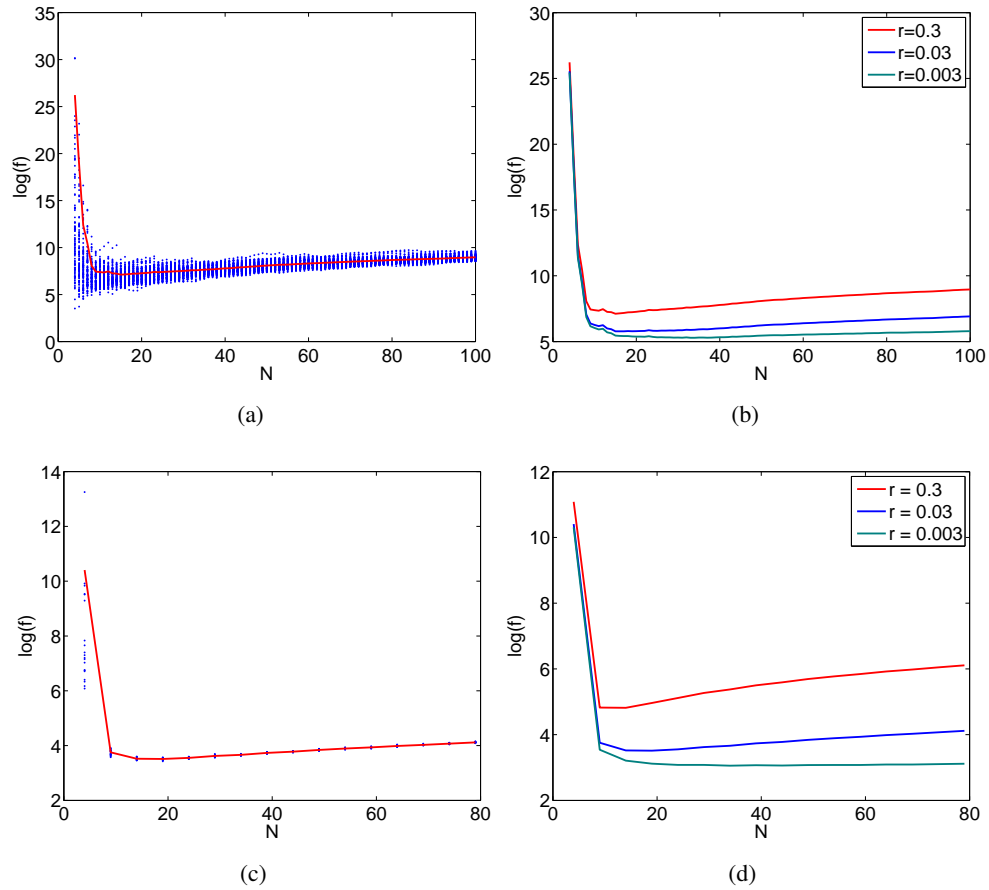


FIGURE 4.1. $\log(f(N))$ en fonction de N ($M = 2$). (a) Résultats pour 100 expériences (bleu) et moyenne (rouge) de la SR 2D translationnelle en fonction de N avec $r = 0.03$. Le minimum est atteint en $N = 12$. (b) Moyenne $\log(f)$ pour différentes valeurs de r pour la SR translationnelle. Minimum à 12 ($r=0.3$), 13 ($r=0.03$) et 30 ($r=0.003$). (c) Super-résolution affine, 20 expériences par valeur de N . Le minimum est atteint en $N = 19$. (d) Moyenne de $\log(f(N))$ pour différentes valeurs de r de la SR affine. Minimum à 14 ($r=0.3$), 19 ($r=0.03$) et 34 ($r=0.003$).

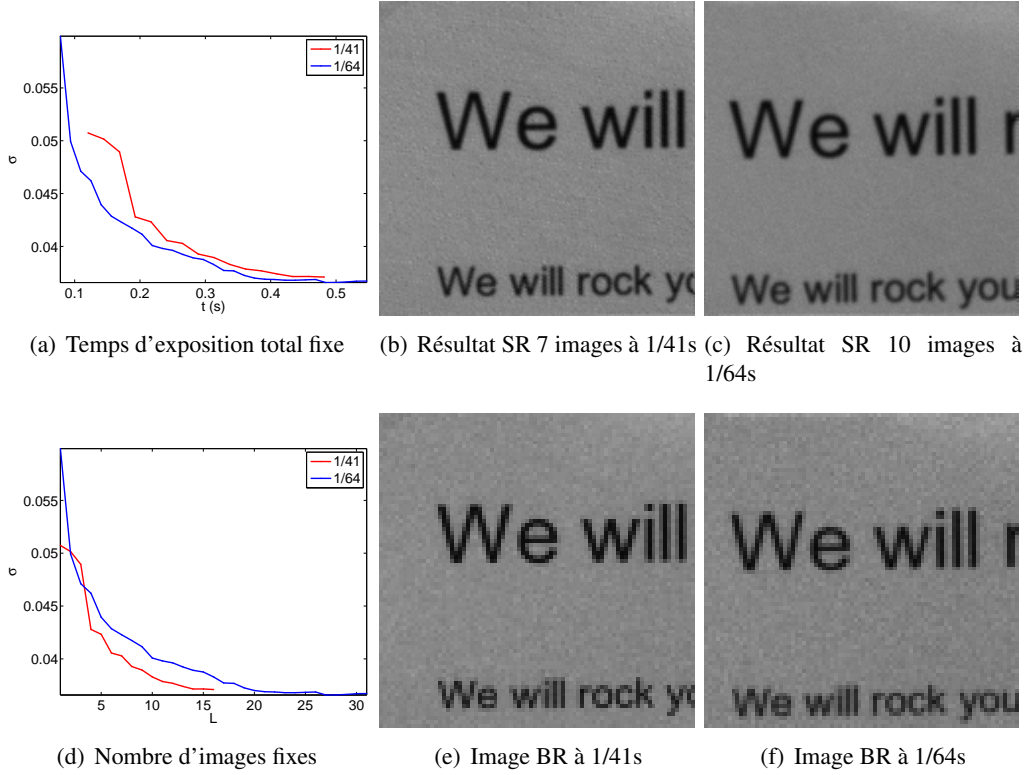


FIGURE 4.2. Erreur de reconstruction en fonction du temps d'exposition total et du nombre d'images ($M = 2$). (a) Erreur en fonction du temps total. (b) Erreur en fonction de N . (c) SR avec 7 images à $1/41$ s. (d) Une image BR à $1/41$ s. (e) SR avec 10 images à $1/64$ s. (f) Une image BR à $1/64$ s.

Ces expériences ont été effectuées pour tester la capacité de notre stratégie d'acquisition à retrouver tout le contenu dynamique et fréquentiel sans régularisation. Les photos ont été prises avec un appareil commercial haut de gamme (photos brutes, format RAW).

Surface plane : On utilise une surface plane illuminée par une source de lumière forte, qui génère 2 niveaux d'irradiance, on est ainsi dans l'hypothèse affine. On calcule 3 temps d'exposition optimaux : $\frac{1}{100}$, $\frac{1}{80}$ et $\frac{1}{32}$ secondes. Ensuite, on prend 20 photos avec un appareil tenu à la main pour chacun de ces 3 temps (le temps total est approximativement de 1.1 secondes). On résout le problème (4.4) avec les 60 images. La figure 4.3 montre des exemples de chaque temps d'exposition et de la saturation correspondante dans le domaine de l'irradiance. On observe l'augmentation de la saturation avec le temps d'exposition.

Dans la figure 4.4, on montre le résultat de la reconstruction HDR-SR de facteur $M = 2$ avec le résultat d'une interpolation bicubique pour pouvoir comparer. Comme attendu, le résultat de la HDR-SR est plus net et moins bruité que l'interpolation bicubique. Dans ce cas, on voit que le contenu fréquentiel n'est pas très riche. Cependant, la reconstruction d'information fréquentielle manquante est visible.

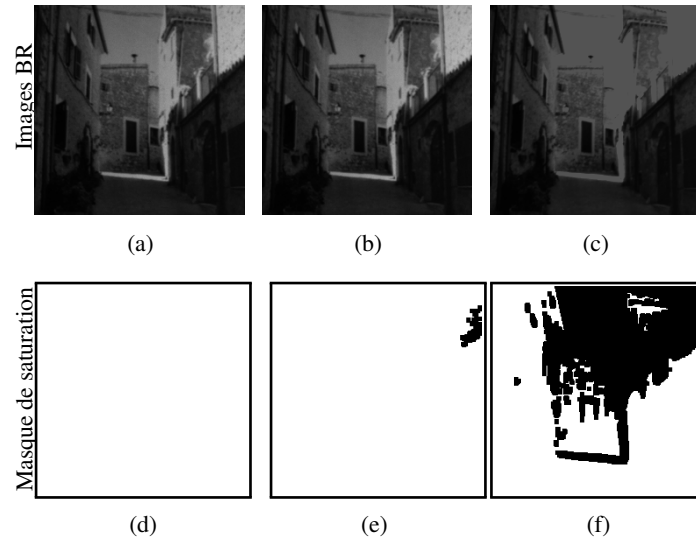


FIGURE 4.3. Images BR et Saturation. Irradiance BR avec différents temps d'exposition et leurs masques de saturation respectifs (parties saturées en noir) (a) image à $\frac{1}{100}$ s, (b) image à $\frac{1}{80}$ s, (c) image à $\frac{1}{30}$ s, (d) saturation pour (a), (e) saturation pour (b), (f) saturation pour (c).

Scène complexe : On fait l'expérience dans une scène plus complexe montrée figure 4.5. On trouve les temps d'exposition suivants : $\frac{1}{64}$, $\frac{1}{41}$ et $\frac{1}{32}$ secondes. On prend 20 images par temps. Le temps total est de 1.4 secondes. On effectue la HDR-SR dans des régions de l'image qui vérifient l'hypothèse affine. Dans la figure 4.6, on montre le résultat de la HDR-SR pour une région sombre et une région claire de la scène. La reconstruction des détails de haute résolution est particulièrement visible sur les parties de l'image en focus (figure 4.6(d) et 4.7). En particulier, le repliement de l'image BR (que l'on remarque sur le texte) est réduit.

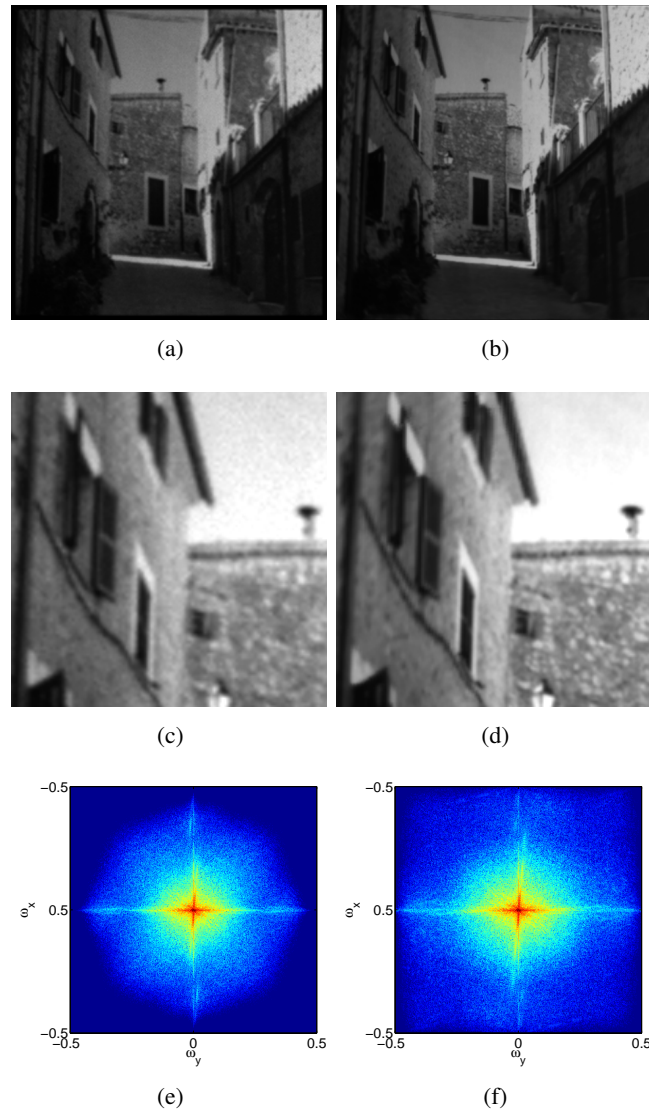


FIGURE 4.4. Résultat de la HDR-SR avec 60 images (a) interpolation bicubique de la première image BR, (b) notre résultat de HDR-SR, (c) détail de (a), (d) détail de (b), (e) Spectre de fréquence 2D de (a), (f) Spectre de fréquence 2D de (b).

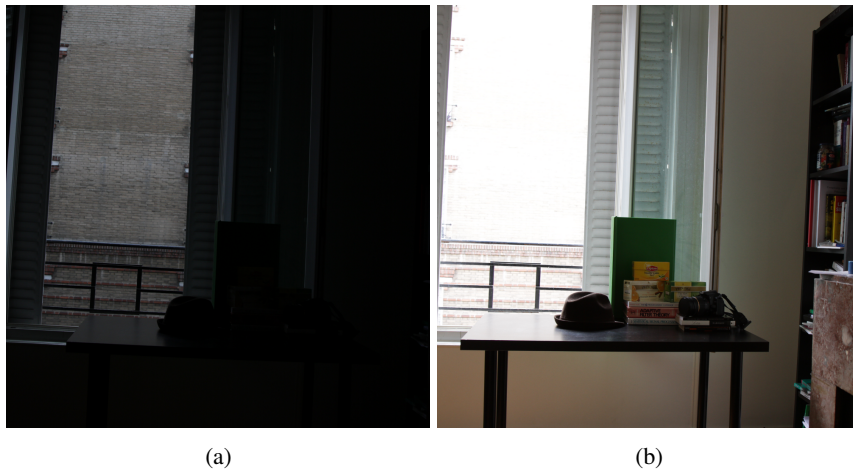


FIGURE 4.5. Scène à 2 temps d'exposition différents.

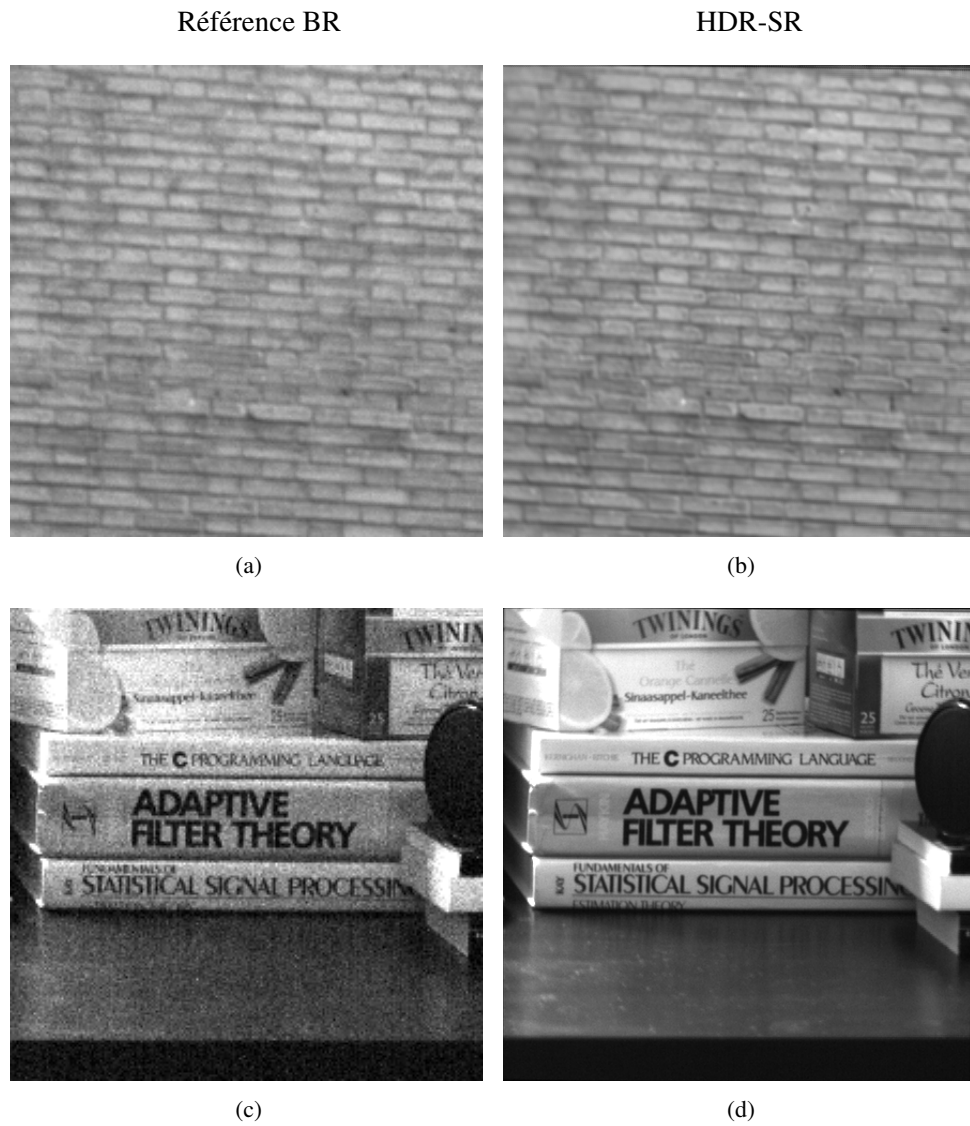


FIGURE 4.6. Résultats de HDR-SR avec 60 images (a) image BR, (b) notre résultat HDR-SR, (c) image BR, (d) notre résultat de HDR-SR

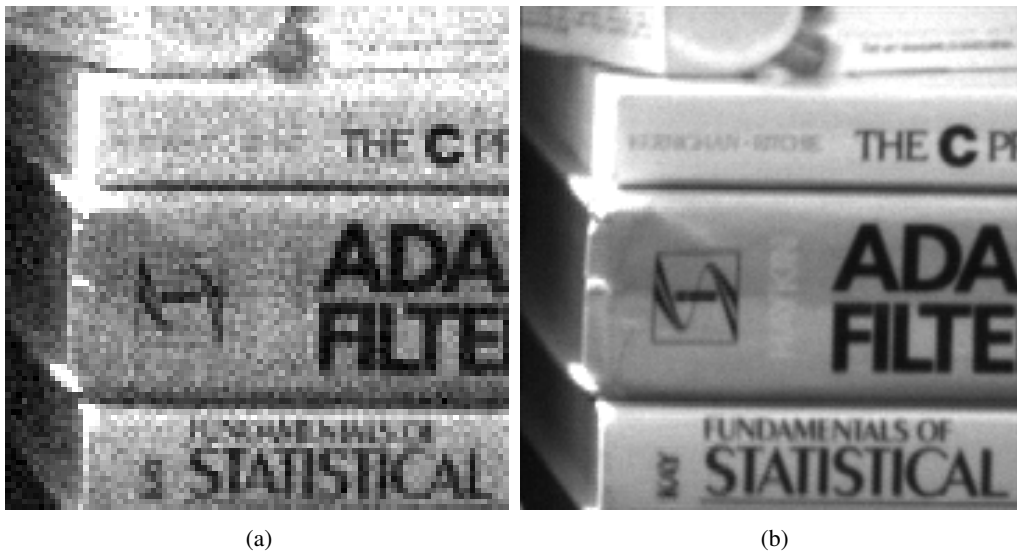


FIGURE 4.7. Détail de la HDR-SR avec 60 images (a) détail de la figure 4.6(c), (b) détail de la figure 4.6 (d).

4.5 Calculs intermédiaires pour la borne sur l'erreur de reconstruction

On utilise d'abord les inégalités de normes d'opérateur

$$\|n_{\text{rec}}\|_2^2 = \|(A'^H W' A')^{-1} A'^H W' n_{\text{in}}\|_2^2 \quad (4.10)$$

$$\leq \|(A'^H W' A')^{-1}\|_{op}^2 \|A'^H W'^{1/2}\|_{op}^2 \|W'^{1/2} n_{\text{in}}\|_2^2 \quad (4.11)$$

$$\leq \frac{\sigma_{\max}^2(A') \sigma_{\max}^2(W'^{1/2})}{\sigma_{\min}^4(A') \sigma_{\min}^4(W'^{1/2})} l^2 N \quad (4.12)$$

$\kappa(N)$ est le conditionnement de $A'^H A'$ (on rappelle que $\kappa(N) = \frac{\sigma_{\max}^2(A')}{\sigma_{\min}^2(A')}$). Comme W' est diagonal,

$$\|n_{\text{rec}}\|_2^2 \leq \frac{\kappa(N)^2}{\sigma_{\max}^2(A')} \frac{(mN/T + \sigma_c^2 N^2/T^2)^2}{\tilde{m}N/T + \sigma_c^2 N^2/T^2} l^2 N \quad (4.13)$$

$$\leq \frac{\kappa(N)^2}{\sigma_{\max}^2(A')} \frac{m^2 l^2 (1 + rN)}{\tilde{m}(1 + \tilde{r}N)} (1 + rN) N^2 \quad (4.14)$$

où $m = \sup(u_0) \asymp \sup(A' u_0)$, $\tilde{m} = \inf(u_0) \asymp \inf(A' u_0)$, $r = \sigma_R^2/(Tm)$, $\tilde{r} = \sigma_R^2/(T\tilde{m})$. En utilisant l'inégalité $\frac{(1+rN)}{(1+\tilde{r}N)} \leq 1$, on a

$$\|n_{\text{rec}}\|_2^2 \leq \frac{\kappa(N)^2}{\sigma_{\max}^2(A')}, \frac{m^2 l^2}{\tilde{m}} (1 + rN) N^2. \quad (4.15)$$

CHAPITRE 5

RÉGULARISATION DES PROBLÈMES DE SUPER-RÉSOLUTION INVERSIBLES MAL CONDITIONNÉS

Lorsque le problème d'interpolation SR par minimisation L^2 est inversible, mais mal conditionné, l'amplification du bruit lors de l'inversion rend inévitable une régularisation ou l'application d'un débruitage après le traitement. On décrit ici le comportement local du conditionnement lorsque l'on suppose des petits mouvements. On en déduit une manière de régulariser qui préserve le signal au mieux possible en faisant intervenir moins de régularisation.

5.1 Conditionnement local

La fusion des grilles BR est une grille d'échantillonnage qui n'est pas périodique en général (exemple figure 5.1). Si le nombre d'images est limité et les mouvements sont petits, on observe des variations locales de la distribution spatiale des échantillons, ce qui mène à une variabilité de la puissance du bruit résiduel après la minimisation L^2 (moindres carrés). Dans cette section, on prédit ce comportement local du bruit à l'aide d'une mesure de conditionnement local dans le cas de petits mouvements (hypothèse justifiée pour le cas d'un appareil photo tenu à la main). On utilise ensuite cette prédiction pour régulariser l'inversion de manière localement adaptative.

On étudie le conditionnement pour des cas proches du cas critique : $N \geq M^2$, N est proche de M^2 et le problème est inversible (voir Chapitre 3). On a observé que lorsque le nombre d'images devient proche de M^2 la probabilité d'avoir un problème mal conditionné est non négligeable (décroissance du conditionnement en $O(\frac{1}{\sqrt{N}})$). On introduit le conditionnement local d'un problème translationnel pur équivalent. On justifie cette notion de conditionnement local en comparant localement le bruit de reconstruction n_R du système et le bruit de reconstruction d'un problème translationnel pur. Lorsque les images BR sont contaminées par un bruit n , on rappelle que le bruit de reconstruction par minimisation L^2 est

$$n_R = A^\dagger n \quad (5.1)$$

On calcule l'énergie de ce bruit localement. On restreint l'espace image de l'application A^\dagger à un pixel BR dans l'espace des images HR pour étudier ce comportement local. Cela revient à restreindre l'espace de départ de l'application A . Cette hypothèse signifie que le bruit de reconstruction est restreint localement. Soit $\mathbf{x}_0 = [x_0, y_0]$ et $\mathbf{x} \in [x_0, x_0 + M - 1] \times [y_0, y_0 + M - 1] = D \subset \mathbb{Z}^2$. Soit $\mathbf{1}_x$ la fonction indicatrice du pixel $\mathbf{x} \in D$ dans les images HR (prend la valeur 1 au pixel \mathbf{x} et 0 sinon).

On considère maintenant la restriction de l'opérateur d'acquisition A à ce domaine D :

$$\begin{aligned} A_{\mathbf{x}_0} : F = \text{vect}((\mathbf{1}_{\mathbf{x}})_{\mathbf{x} \in D}) &\rightarrow E = A(\text{vect}((\mathbf{1}_{\mathbf{x}})_{\mathbf{x} \in D})) \\ u &\rightarrow A_{\mathbf{x}_0} u \end{aligned} \quad (5.2)$$

On appelle conditionnement local $\kappa(\mathbf{x}_0)$ à la position \mathbf{x}_0 , le conditionnement de $A_{\mathbf{x}_0}$. Le conditionnement est le ratio des bornes supérieure et inférieure de la quantité (plus grande et plus petite valeur singulière)

$$\|A_{\mathbf{x}_0} u\|, \|u\| = 1. \quad (5.3)$$

On étudie le comportement de cette quantité. Pour cela, on considère une image $u \in F$ décomposée dans la base des pixels HR $(\mathbf{1}_{\mathbf{x}_k})_k$. Soit $u = \sum_k b_k \mathbf{1}_{\mathbf{x}_k} \in F$ avec $\|u\| = 1$. On a :

$$\|A_{\mathbf{x}_0} u\|^2 = \left\| \sum_k b_k A \mathbf{1}_{\mathbf{x}_k} \right\|^2 \quad (5.4)$$

$$= \sum_{k_1, k_2} \bar{b}_{k_1} b_{k_2} (\mathbf{1}_{\mathbf{x}_{k_1}})^H A^H A \mathbf{1}_{\mathbf{x}_{k_2}} \quad (5.5)$$

$$= \sum_{k_1, k_2} \bar{b}_{k_1} b_{k_2} \sum_{i=1, N} \sum_{\mathbf{y} \in \Gamma} \text{sincd}(\mathbf{y} - \tau_{i, k_1}) \text{sincd}(\mathbf{y} - \tau_{i, k_2}) \quad (5.6)$$

où sincd est l'interpolateur de Shannon fini discret. Les vecteurs $\tau_{i, k} = q_i \mathbf{x}_k$ sont les translations induites par les mouvements affines et Γ est la grille d'échantillonnage HR. Comme la fonction sincd est différentiable, on peut utiliser le théorème des accroissements finis pour comparer cette expression à une expression translationnelle pure et obtenir :

$$\left| \|A_{\mathbf{x}_0} u\|^2 - \|A_{\mathbf{x}_0}^{tr} u\|^2 \right| \leq c \|\boldsymbol{\theta} - \boldsymbol{\theta}^{tr}\|^2 \quad (5.7)$$

où $\boldsymbol{\theta} = (\tau_{i, k})_{i, k}$ est l'ensemble des translations induites par le mouvement. c est une constante qui dépend du gradient de la fonction sincd . Le vecteur $\boldsymbol{\theta}^{tr}$ est $\boldsymbol{\theta}$ moyenné sur le pixel BR (sur l'indice k , décrit le domaine D) et $A_{\mathbf{x}_0}^{tr}$ est l'opérateur translationnel SR pur associé à $\boldsymbol{\theta}^{tr}$. La constante c ne dépend pas de \mathbf{x}_0 . On déduit que, pour des mouvements suffisamment petits, l'énergie du bruit du système se comporte de la même façon que dans le cas translationnel. Les expériences montrent que pour des affinités dans une petite plage de valeur (rotations entre -5 et $+5$ degrés, zoom entre $\times 0.95$ et $\times 1.05$), on peut utiliser $\kappa(\mathbf{x}_0) = \text{cond}(R)$ comme une mesure locale du conditionnement, avec R défini comme dans la section 3.2 avec les translations $\boldsymbol{\theta}^{tr}$. Le résultat de cette prédiction est montré dans la figure 5.1. La carte de prédiction du conditionnement local correspond bien à la carte d'intensité du bruit de reconstruction. On observe par ailleurs que les variations du bruit (et du conditionnement) peuvent être très forte à l'intérieure d'une image HR reconstruite par minimisation L^2 .

Dans la figure 5.2, on étudie expérimentalement la plage de validité de l'estimation du conditionnement local. Pour chaque expérience, on génère des translations aléatoires dans $[0, M]$. On fait varier la plage des paramètres aléatoires de zoom et de rotation. Lorsque cette plage augmente, les variations de l'estimateur de conditionnement local deviennent trop grandes pour pouvoir être utilisées comme prédiction de la quantité de régularisation. On étudie, figure 5.3, la validité de l'hypothèse lorsque la plage de rotation augmente. La corrélation entre le conditionnement local et le bruit de reconstruction observé diminue.

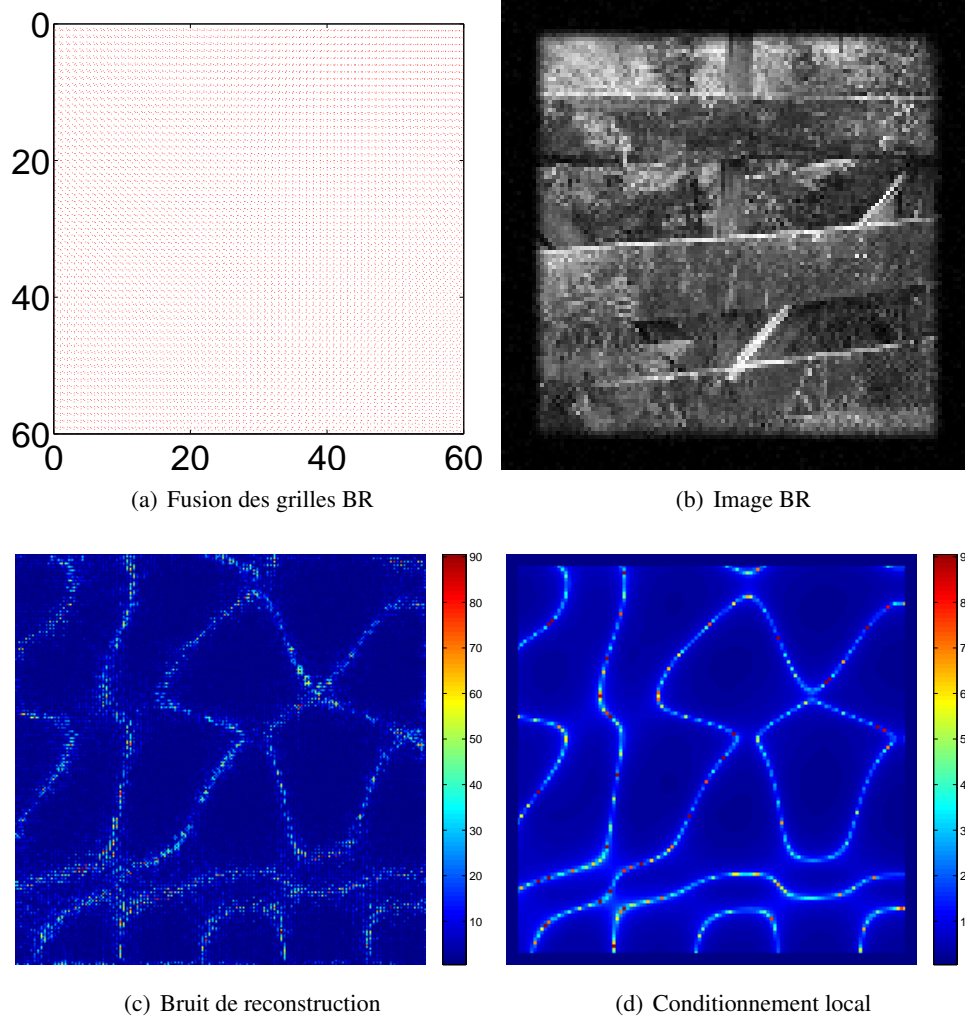


FIGURE 5.1. Conditionnement local du problème de SR. (a) Zoom sur la fusion des 4 grilles BR (60×60 pixels coin en haut à gauche). (b) Exemple d'image BR. (c) Amplitude du bruit de reconstruction (échantillonné sur la grille BR) normalisé par la variance du bruit d'entrée. (d) Conditionnement local $2\sqrt{\kappa(\mathbf{x})}$.

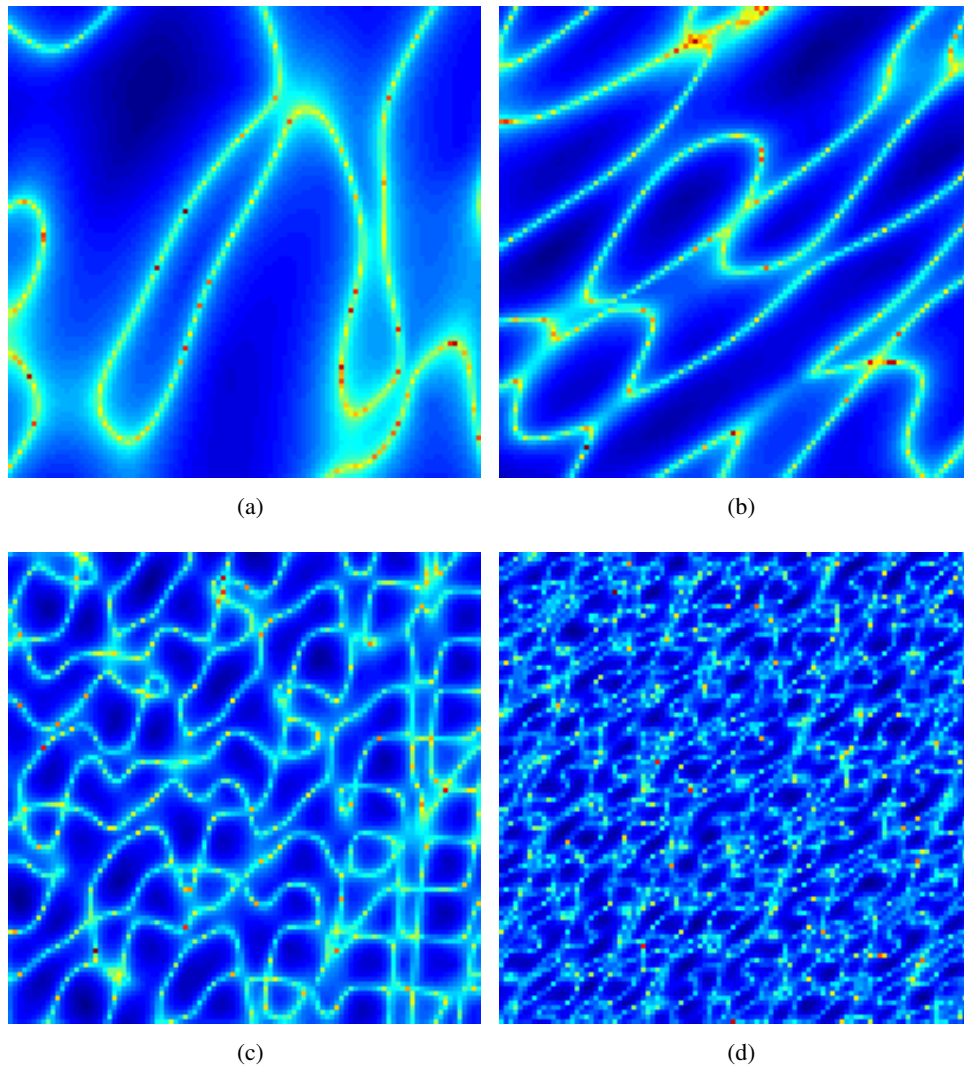


FIGURE 5.2. Variabilité de l'approximation de conditionnement local. Carte de la prédiction $\log(\kappa_x)$ pour différentes variabilités des paramètres. (a) Zoom entre -1% et 1%, rotations entre -1 et 1 degrés, (b) Zoom entre -2.5% et 2.58%, rotations entre -3 et 3 degrés. (c) Zoom entre -5% et 5%, rotations entre -5 et 5 degrés. (d) Zoom entre -10% et 10%, rotations entre -20 et 20 degrés.

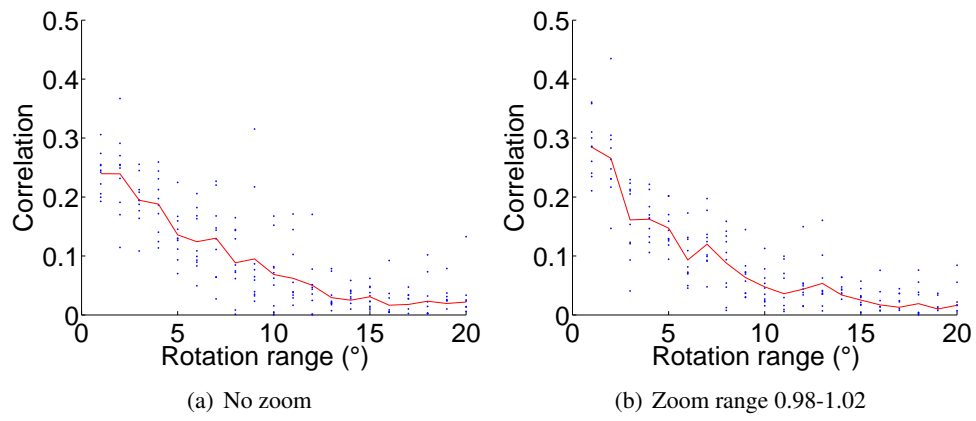


FIGURE 5.3. Validité de l'approximation de conditionnement local. Corrélacion entre $\log(\kappa)$ et $\log(n_R^2)$. En bleu, les expériences, en rouge la moyenne.

5.2 Régularisation locale

Pour la super-résolution, deux types de régularisation ont été principalement utilisés : la régularisation de Tychonoff [53] et la régularisation TV (ou TV anisotrope) [43, 45, 57, 113]. On observe le comportement local de la régularisation induite par la troncature du gradient. On propose ensuite une régularisation locale par variation totale où notre mesure de conditionnement local définit des poids pour le terme de variation totale.

5.2.1 Régularisation locale par troncature du gradient conjugué

L'algorithme du gradient conjugué (CG pour "conjugate gradient") est une manière rapide de résoudre un problème de minimisation L^2 [49] lorsque le système est bien conditionné. Dans le cas contraire, on a observé que le gradient conjugué tronqué régularise principalement dans les zones mal conditionnées. Lorsque l'on effectue le gradient conjugué, on calcule l'application linéaire inverse de manière itérative. On rappelle qu'un système utilisant une approximation de rang r de l'identité est résolue par au plus $r + 1$ itérations de CG (la démonstration peut être trouvée dans [49]). Tronquer le CG revient à résoudre le problème avec une régularisation de Tychonoff sur les plus petites valeurs propres (pas montré dans [49]) :

Proposition 5.2.1. *Soit \tilde{u}_r le résultat de l'algorithme de CG tronqué à r itérations pour le problème $\operatorname{argmin}_u \|Au - w\|_2^2$. Soit :*

$$u_r = \operatorname{argmin} \|A(I - P)u - w\|^2 + \epsilon \|Pu\|_2^2 \quad (5.8)$$

avec ϵ la $r + 1$ plus grosse valeur propre de $A^H A$ et P un projecteur sur les espaces propres ayant des valeurs propres plus petites que ϵ , alors :

$$\|\tilde{u}_r - u_r\|_2 \leq c\epsilon \quad (5.9)$$

où c est une constante qui dépend de r .

Démonstration. On décompose $A^H A$ comme une approximation de rang r de l'identité plus une matrice résiduelle :

$$A^H A = \epsilon.I + D + L$$

où ϵ est la $r + 1$ -ème plus grande valeur propre, D est une matrice de rang r obtenue en soustrayant ϵ des r plus grandes valeurs propres. La matrice L résultante a une norme de Frobenius $\|L\|_F < \epsilon$. Étant donné une inconnue, trouver la solution de la r -ième itération de CG est une application lipschitzienne par rapport à la norme de Frobenius de la matrice à inverser. On a donc que le résultat \tilde{u}_r de la r -ème itération avec $A^H A$:

$$\|\tilde{u}_r - u_r\| \leq c\epsilon$$

où c est la constante de Lipschitz et u_r est le résultat de la r -ème itération avec $\epsilon.I + D$, ce qui correspond à la minimisation de l'équation (5.8) \square

Cette approximation est meilleure lorsque ϵ est petit. La troncature du gradient conjugué revient à une régularisation de Tychonoff qui cherche à minimiser la norme sur les plus petites valeurs propres de $A^H A$. Elle n'est donc pas particulièrement conçue pour des

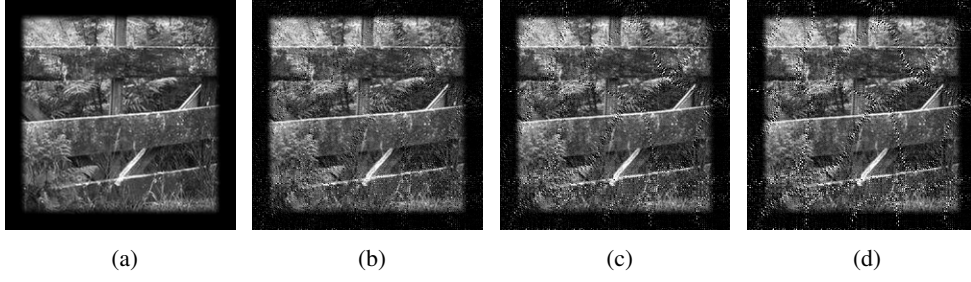


FIGURE 5.4. Influence du nombre d'itérations de gradient conjugué sur la solution. (a) image HR idéale (b) Résultat après 50 itérations (c) après 100 itérations et (d) 200 itérations.

images naturelles, car elle dépend uniquement de l'opérateur d'acquisition A et non d'une hypothèse de régularité sur le contenu de la scène à photographier.

Les espaces propres associés à ces petites valeurs propres sont principalement localisés dans les zones mal conditionnées. Cela vient du fait que la valeur propre la plus grande est associée aux images constantes, et que les valeurs propres sont positives. Comme la décomposition sur une base propre préserve l'énergie, l'énergie associée aux plus petites valeurs propres est localisée dans les zones mal conditionnées. On peut donc considérer que le CG tronqué est une régularisation locale. Ce fait est illustré dans la figure 5.4, avec l'effet de la troncature à différentes étapes. Dans cette expérience, le bruit de reconstruction est minimisé pour une certaine troncature du gradient conjugué. Lorsque le nombre d'itérations croît, on observe que l'erreur a un comportement local comme on s'y attend, car le gradient conjugué tend vers la solution de la minimisation L^2 .

5.2.2 Régularisation locale par pondération de la TV

Dans [97], les auteurs proposent de pondérer la variation totale anisotrope par les termes diagonaux de l'opérateur ce qui est une façon empirique de prendre en compte le conditionnement local. On cherche à minimiser l'attache aux données L^2 régularisée par variation totale :

$$\begin{aligned} J_\alpha(u) &= \|Au - w\|_2^2 + \lambda TV_\alpha(u) \\ TV_\alpha(u) &= \int \alpha \cdot |\nabla u| \end{aligned} \quad (5.10)$$

où $\alpha(\mathbf{x}) = \min(r, \log(\kappa))$ avec r un paramètre de tolérance et λ est le paramètre de régularisation global, que l'on suppose connu et fixe. Lorsque α est constant, TV_α est la régularisation par variation totale classique. Dans [106], on a d'abord choisi $\alpha(\mathbf{x}) = \log(\kappa(\mathbf{x}))$ expérimentalement. On justifie qualitativement ce choix plus fin pour α en utilisant un modèle simple. Dans la pratique, on utilise une version discrétisée lissée de la fonction TV .

Pour expliquer notre choix de la fonction α , on considère un problème translationnel pur avec régularisation TV globale. On cherche à savoir de quelle manière le paramètre de régularisation doit être réglé pour que le contenu fréquentiel de l'image HR soit respecté au maximum. La régularisation peut échouer et détruire des détails de l'image. En revanche, le bruit généré par l'inversion n'est pas corrélé à l'image. Notre objectif est de préserver au

maximum les données. On cherche donc un α_0 qui minimise la fonction de risque suivante (avec ϵ^2 un paramètre de tolérance) sous un modèle gaussien pour n :

$$\begin{aligned} \alpha_0 &= \operatorname{argmin}_{\alpha} \min(E\|\tilde{u}_{\alpha} - u\|^2, \epsilon^2) \\ \text{avec } \tilde{u}_{\alpha} &= \operatorname{argmin}_{\tilde{u}} J_{\alpha}(\tilde{u}). \end{aligned} \quad (5.11)$$

Autrement dit, on préfère un petit bruit de reconstruction d'énergie ϵ^2 à une régularisation donnant la même erreur.

Il a été montré dans [27] que la solution d'un problème avec régularisation TV peut être calculée grâce à un algorithme de point fixe utilisant une linéarisation du gradient de la variation totale $\partial TV(u)$. Cet algorithme est aussi appelé méthode de Weiszfeld. C'est l'algorithme que nous avons utilisé pour les expériences à la fin de cette section. Il permet aussi d'avoir un modèle linéaire de la fonctionnelle autour de la solution (lorsque l'algorithme a convergé). À chaque itération n , on linéarise ce terme en utilisant une estimation de l'image HR précédente \tilde{u}_{n-1} : $(\partial TV)_n \approx B_n = -\operatorname{div} \frac{\nabla}{|\nabla \tilde{u}_{n-1}|}$. Soit B cet opérateur linéaire, lorsque l'algorithme a convergé ($\tilde{u}_n \approx u_0$). On a (on rappelle que $R = A^H A$) :

$$\tilde{u}_{\alpha} = (R + \alpha \lambda B)^{-1} A^H w. \quad (5.12)$$

En utilisant l'indépendance entre signal et bruit,

$$\begin{aligned} E\|\tilde{u}_{\alpha} - u_0\|^2 &= \|(R + \alpha \lambda B)^{-1} \alpha \lambda B u_0\|^2 \\ &+ E\|(R + \alpha \lambda B)^{-1} A^H n\|^2. \end{aligned} \quad (5.13)$$

On écrit cette égalité dans chaque sous-espace propre de $R + \alpha \lambda B$. Si R et B commutent, les valeurs propres sont la somme des valeurs propres de R et B . On suppose qu'elles commutent approximativement. Soit σ_i^2 les v.p. de R et μ_i les v.p. associées de λB dans la diagonalisation jointe. Alors on peut décomposer le risque $\rho(\alpha) = E\|\tilde{u}_{\alpha} - u_0\|^2 = \sum_i \rho_i^2$ avec :

$$\rho_i^2 \approx (\sigma_i^2 + \alpha \mu_i)^{-2} (\alpha^2 \|\lambda(Bu_0)_i\|^2 + \|(A^H n)_i\|^2). \quad (5.14)$$

Chaque ρ_i^2 est minimisé par $\alpha_0 = \frac{\mu_i \|(A^H n)_i\|^2}{\sigma_i^2 \|\lambda(Bu_0)_i\|^2} = \frac{\|n_i\|^2}{\mu_i \|u_i\|^2}$ (proche de la solution du filtrage de Wiener). Si la régularisation est parfaite (c'est-à-dire $\mu_i \|u_i\|^2$ est une constante), le paramètre de régularisation ne dépend pas de A . Cependant, le comportement de $\sum_i \rho_i^2$ sous un modèle de bruit pour $\mu_i \|u_i\|^2$ est en général plus complexe. Si les deux matrices ne commutent pas les v.p. de $R + \alpha \lambda B$ sont bornées par $\sigma_i^2 + \mu_{max}$ et on peut borner le risque dans les espaces propres grâce à cette inégalité.

En réécrivant l'équation 5.11 dans ces espaces propres, on obtient

$$\alpha_0 = \operatorname{argmin}_{\alpha} \min \left(\sum_i \rho_i^2, \epsilon^2 \right) \quad (5.15)$$

On représente figure 5.5 la relation entre le α optimal et $\log(\kappa)$. Pour vérifier ce résultat, une expérience a été générée en calculant le α optimal pour différentes réalisations d'une SR translationnelle avec régularisation TV pour l'image *baboon*. L'autre est générée en utilisant la minimisation (5.15) et des paramètres constants pour μ_i , $|(Bu)_i|$ et $|n_i|$. Dans

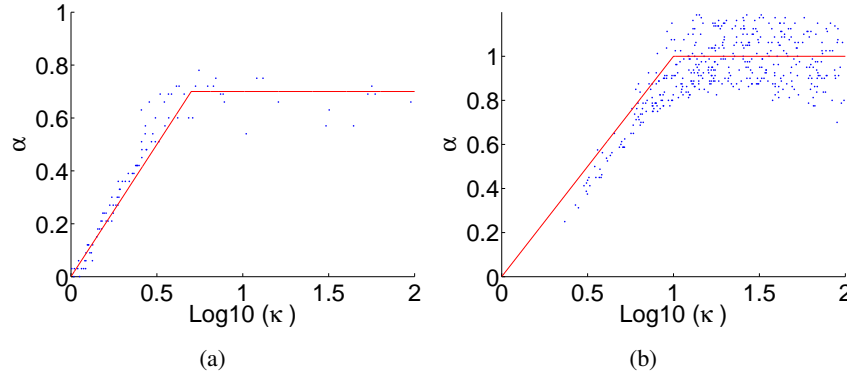


FIGURE 5.5. Paramètre optimal de régularisation en fonction du conditionnement. Notre choix pour α est la courbe en rouge. (a) Paramètre de régularisation optimal en fonction du conditionnement pour la SR translationnelle. (b) Paramètre de régularisation pour la SR translationnelle obtenu en minimisant la fonction de risque de l'équation (5.15).

les deux cas, on observe que ce α optimal suit le modèle de fonction du conditionnement que l'on a choisi.

La complexité de notre algorithme de régularisation locale est la complexité de la super-résolution TV globale plus le coût du calcul du conditionnement local. Ce coût est $O(C_{M,N}(ML)^2)$, où $C_{M,N}$ est le coût du calcul du conditionnement d'un système $M \times N$ (les systèmes SR sont calculés rapidement dans le domaine de Fourier comme montré dans l'équation (3.22)). Le conditionnement local doit être utilisé lorsque le nombre d'images est proche de M^2 . Le coût additionnel de ce calcul est petit comparé au coût de la SR régularisée par TV qui est de $O(K_{it}C_{SR})$ opérations, où K_{it} est le petit nombre d'itérations (typiquement 3) requis pour la convergence de l'algorithme de point fixe, et C_{SR} est le coût de la minimisation SR L^2 sans régularisation.

5.3 Expériences

On effectue une première expérience de super-résolution synthétique avec $M = 2$. On génère $N = 4$ images BR bruitées à partir d'une image HR de taille 240×240 (SR avec $M = 2$, rotations entre -5 et 5 degrés, translations uniformément distribuées dans $[0, M]^2$, zoom entre $0.95, 1.05$). L'image HR ainsi qu'une image BR sont représentées figure 5.6. On montre figure 5.8 le résultat de la régularisation locale, comparé à une régularisation globale. On effectue l'interpolation SR sans régularisation, avec régularisation TV globale et locale avec paramètre optimal. On a utilisé notre connaissance de l'image parfaite pour déterminer le paramètre de régularisation donnant le meilleur PSNR pour les deux types de régularisation. Pour la régularisation globale, l'image résultante est excessivement lissée dans les endroits ayant un meilleur conditionnement, et pas assez à d'autres endroits. Avec la régularisation locale, il y a lissage seulement dans les zones mal conditionnées. Cela donne une meilleure reconstruction de l'image HR. Les différences sont particulièrement visibles sur les images représentant le bruit de reconstruction ainsi que le montre l'agrandissement de la figure 5.7.

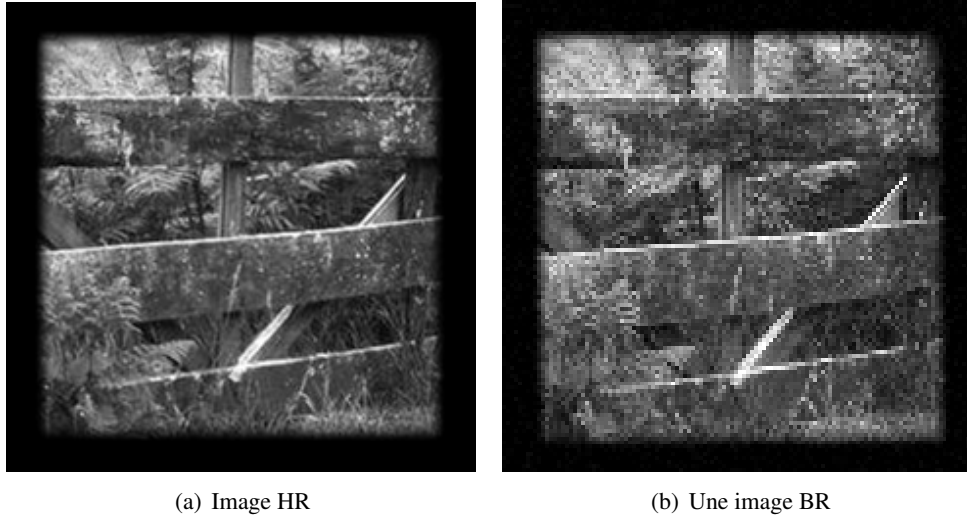


FIGURE 5.6. Images HR et BR de la première expérience.

Dans la figure 5.9, on effectue une autre expérience synthétique avec $M = 3$. On génère cette expérience de la même façon que la précédente avec $N = 10$. La préservation du signal est améliorée avec la régularisation locale par rapport à la régularisation globale. On observe mieux cette préservation sur les cartes de différences (figure 5.9 (e)(f)).

figure 5.10, notre régularisation locale est utilisée pour des images réelles. On utilise 4 images de l'expérience 2 de la figure figure 3.9 (Section 3.2). Le résultat de la SR sans régularisation montre le comportement local du bruit décrit précédemment. Ce comportement est bien prédit par le conditionnement local. Le résultat de la SR avec régularisation locale a amélioré la netteté dans les zones bien conditionnées par rapport à la régularisation globale. La zone qui bénéficie de cette régularisation est montrée figure 5.11.

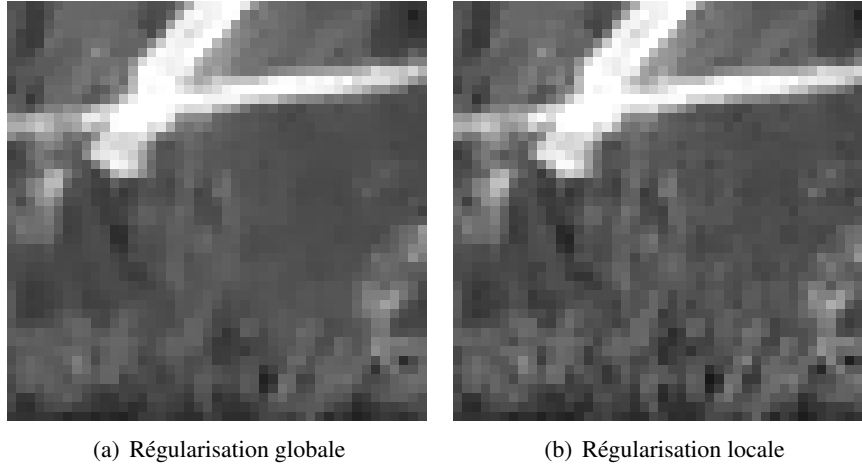


FIGURE 5.7. Détail régularisation TV locale. Détails de la figure 5.8.(e) et 5.8.(g).

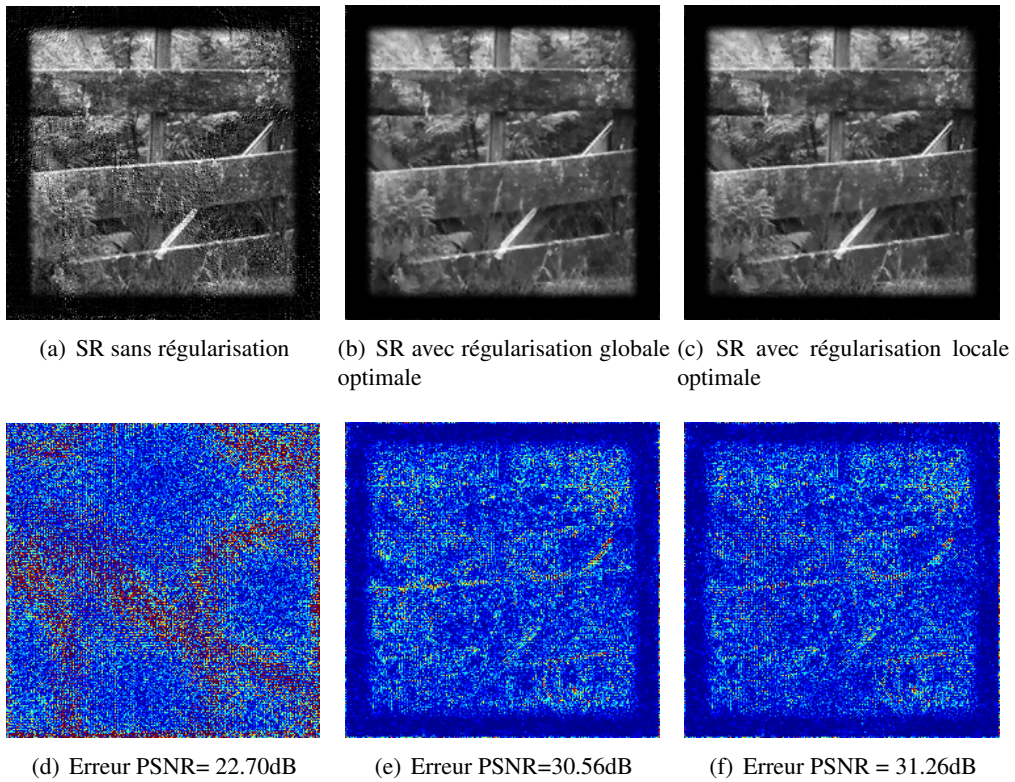
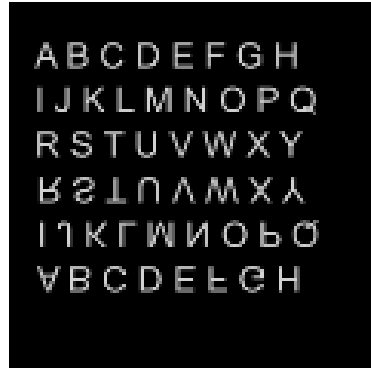
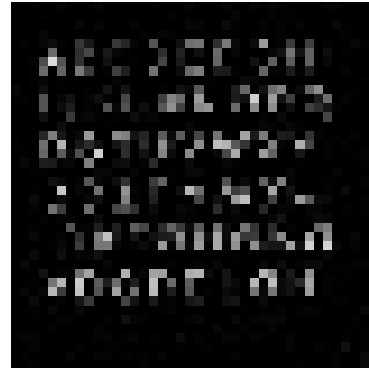


FIGURE 5.8. Régularisation TV locale pour une super-résolution critique $M = 2$. Erreur de reconstruction dans une échelle bleu rouge représentant les niveaux de gris de l'image $[0,30]$ (les images sont dans $[0,255]$).



(a) Image HR



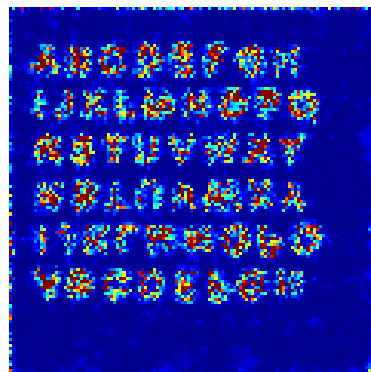
(b) Une image BR



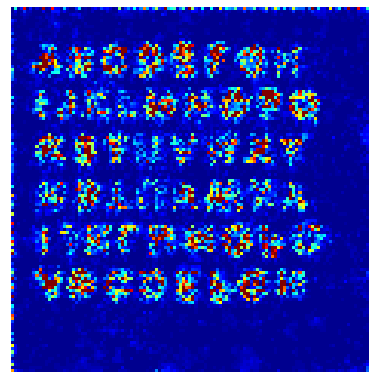
(c) SR avec régularisation globale PSNR=23.98dB



(d) SR avec régularisation locale PSNR=25.09dB



(e) Erreur de reconstruction(c)-(a)



(f) (d)-(a)

FIGURE 5.9. Régularisation TV locale pour une super-résolution critique $M = 3$. Erreur de reconstruction dans une échelle bleu rouge représentant les niveaux de gris de l'image [0,30] (les images sont dans [0,255]).

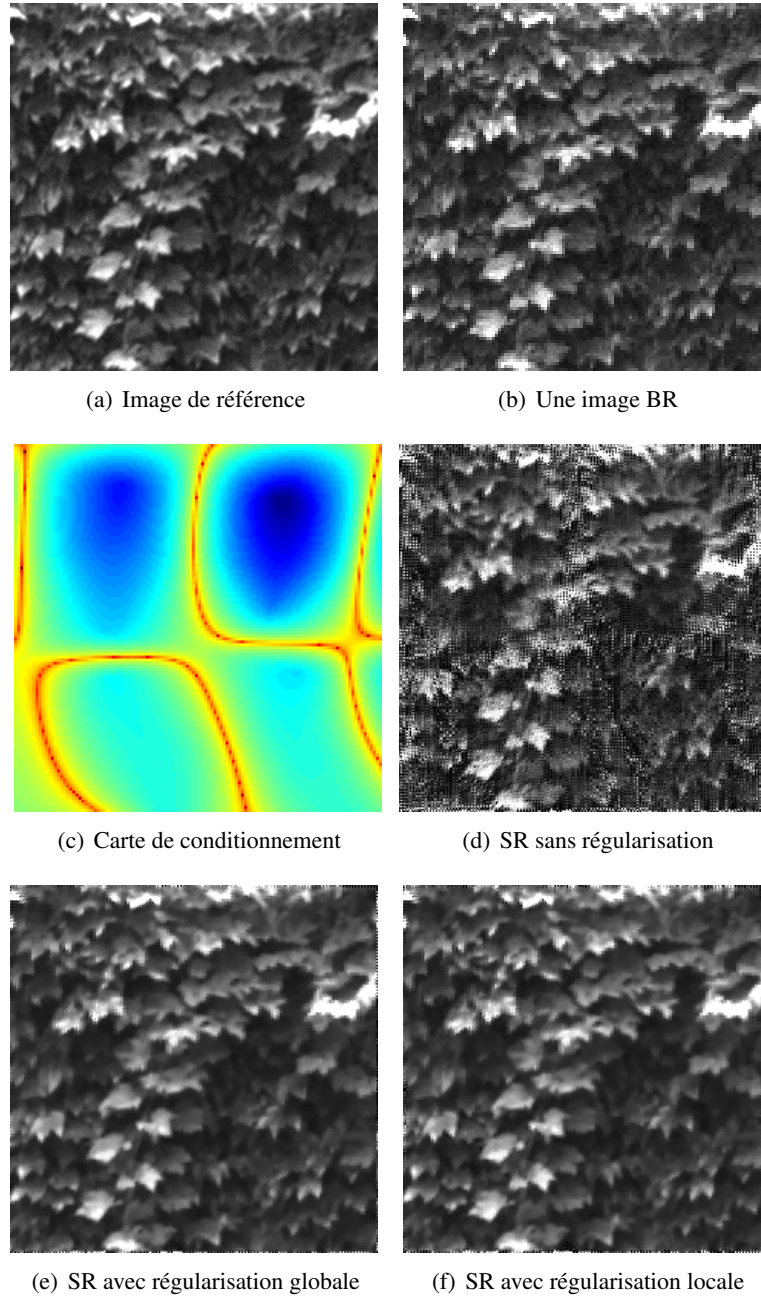


FIGURE 5.10. Régularisation TV locale avec des images réelles $M = 2$. Comparaison entre régularisation globale et locale avec 4 images réelles.

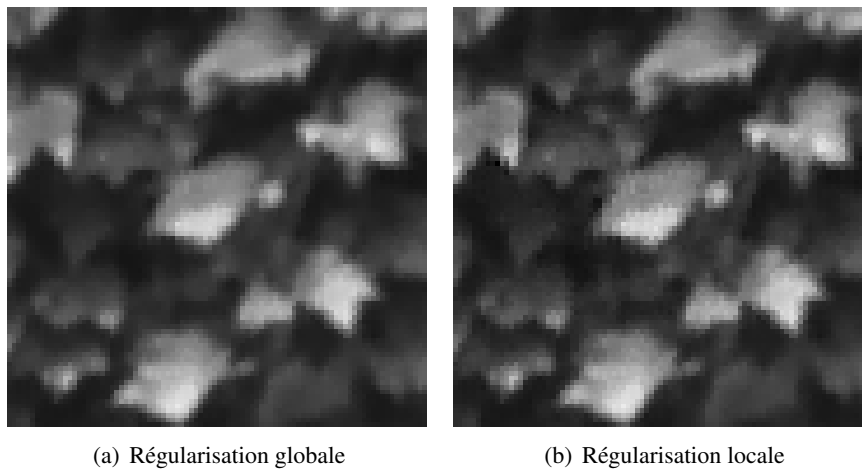


FIGURE 5.11. Détail régularisation TV locale. Détails de la figure [5.10.\(e\)](#) et [5.10.\(f\)](#)

5.3.1 Application au dématricage multi-image

On montre ici que l'on peut appliquer notre algorithme pour effectuer un dématricage. En général, les capteurs pour les différentes couleurs sont entrelacés pour former une matrice de capteurs. La disposition de ces capteurs est la matrice de Bayer. Les composantes rouge vert bleu (RVB) sont entrelacées de telle manière que dans un carré formé de 4 pixels, il s'y trouve un de rouge, un de bleu et deux de vert. De nombreux algorithmes performants permettent de produire une image ayant de l'information RVB pour chaque pixel. On appelle cette opération le dématricage. [42] a montré les bénéfices du dématricage multi-image. Parce qu'il n'y a aucune hypothèse faite sur la régularité de l'image, reconstruire une image HR est possible lorsque le dématricage mono-image échoue. Si l'utilisateur a l'opportunité de prendre 4 photographies de la même scène, on peut utiliser notre régularisation locale pour reconstruire l'image HR avec $M = 2$ avec les composantes RVB indépendamment. De cette manière, on limite la quantité de régularisation utilisée pour produire l'image finale. On montre figure 5.12 la comparaison entre un dématricage multi-image avec bruit et un dématricage mono-image de l'état de l'art (dématricage par autosimilarité [8]) avec bruit. On a généré un exemple synthétique en générant 4 versions BR d'une image HR et en ajoutant du bruit. L'entrée pour le dématricage mono-image est la matrice de Bayer générée à partir de l'image HR. On a spécifiquement choisi une image où le dématricage mono-image échoue pour illustrer les bénéfices possibles de la super-résolution avec régularisation locale. Le dématricage multi-image avec notre régularisation TV locale produit une image HR sans anomalie chromatique, ce qui n'est pas le cas du dématricage mono-image. On voit donc que même dans le cas mal conditionné, on a réussi à produire une image HR de bonne qualité en limitant autant que possible la régularisation.

5.4 Conclusion

Ce chapitre a permis de déterminer avec précision le comportement du bruit de reconstruction d'une minimisation L^2 . On a exploité cette information, pour minimiser la quantité de régularisation à effectuer. Idéalement, le paramètre de tolérance, utilisé pour favoriser le bruit non corrélé avec le signal (quitte à ce qu'il soit plus fort), devrait être déterminé en accord avec une méthode de débruitage à effectuer après la super-résolution. Par exemple, une méthode (Non Local Means) à base d'imagettes (patches) utilisant l'autosimilarité des images naturelles pourrait agir dans les zones mal conditionnées, et utiliser la connaissance de la quantité de bruit dans chaque imagette pour éviter de les utiliser pour le débruitage. Dans le cas de bruits à variance variable, il a été montré [3] qu'une telle adaptation améliore les performances de débruitage. Les évolutions récentes des algorithmes non-locaux [62] profiteraient de la même manière d'une telle connaissance. Une fois déterminée la performance de tels algorithmes pour un niveau de bruit donné, on pourrait régler la tolérance pour accepter un niveau de bruit que l'on est sûr de pouvoir enlever après.

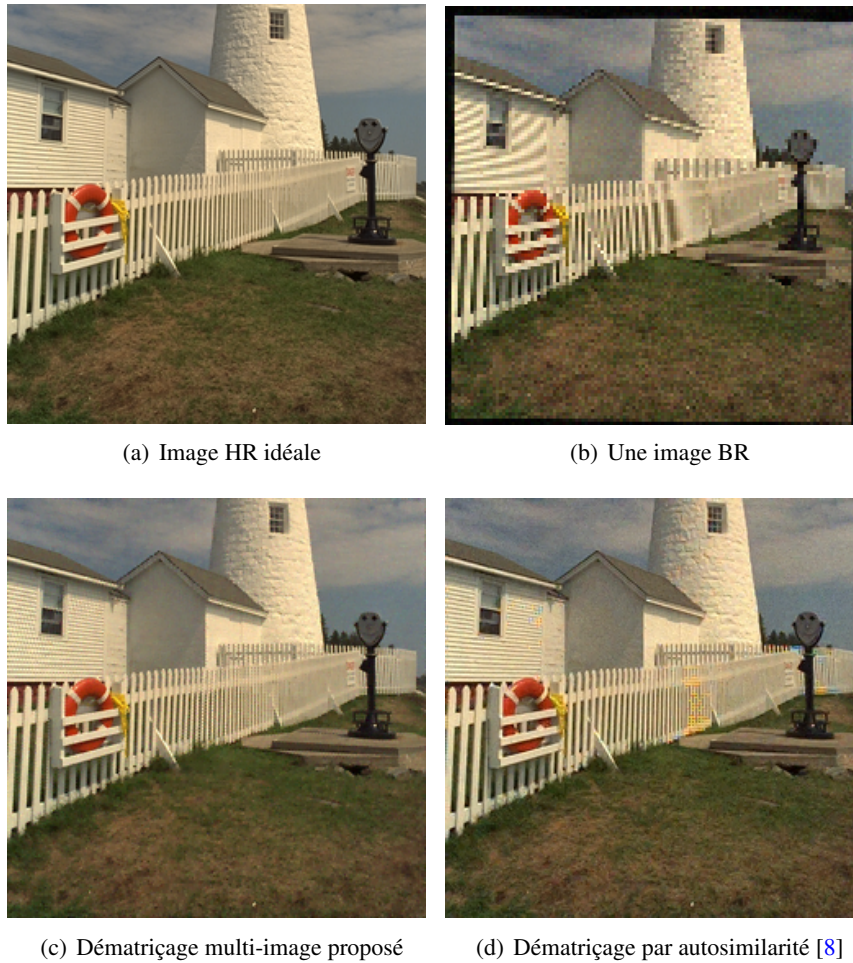


FIGURE 5.12. Dématriage multi-image avec 4 images BR avec régularisation TV locale. Les images BR sont simulées à partir de l'image idéale HR en sous-échantillonnant chaque composante RVB par un facteur $M = 2$ après mouvement aléatoire, un bruit gaussien de déviation standard $\sigma = 5$ est ajouté (Pixels ont des valeurs dans $[0, 255]$).

CHAPITRE 6

ROBUSTESSE DE LA SUPER-RÉSOLUTION L^1 AUX DONNÉES ABERRANTES

La norme L^2 a fréquemment été choisie pour quantifier l'attache aux données car elle correspond aux propriétés statistiques du bruit d'acquisition. Une capacité de réduction du bruit découle de ces propriétés (Chapitre 3). Plus récemment, la minimisation L^1 a été utilisée pour débruiter des images contaminées par des données aberrantes (outliers) [77]. Il a été montré qu'elle permettait une robustesse aux outliers dans le cadre de la super-résolution [45, 113]. Cependant, une régularisation est généralement ajoutée à l'attache aux données, ce qui revient à supposer qu'il n'y a pas assez d'information pour retrouver l'image HR. Dans le chapitre 3, on montre que lorsque le nombre d'images BR est suffisant, il est probable que la régularisation devienne inutile, mais ces résultats s'appliquent mal au cas des données aberrantes (données corrompues, outliers) que l'on peut considérer comme un bruit non borné en amplitude a priori.

Dans ce chapitre, on détermine des conditions pour lesquelles l'interpolation de super-résolution (mouvements supposés connus) est bien informée dans le cas d'un bruit d'outliers. On montre que c'est le cas (en probabilité), si le pourcentage d'images contaminées est inférieur à une constante. On montre par ailleurs que cette borne linéaire est précise (au sens de son évolution en fonction du nombre d'images).

6.1 Lien entre robustesse et reconstruction parcimonieuse

6.1.1 État de l'art

Candès et Tao [24] ont montré que la robustesse aux outliers par minimisation L^1 (dans le cas d'observations redondantes) est équivalente à un problème de reconstruction parcimonieuse. Ils en déduisent une condition de reconstruction parfaite (et approchée dans le cas bruité) basée sur une propriété d'isométrie approchée de la matrice d'observation du problème équivalent, la *Restricted Isometry Property* (RIP). Pour des matrices d'observation aléatoires, ils déduisent que la minimisation L^1 est robuste à un nombre d'outliers qui suit une loi asymptotique meilleure que linéaire en fonction du nombre d'observations : pour des observations comportant des erreurs, le pourcentage d'erreur qu'il est possible de corriger augmente avec le nombre d'observations.

Il faut noter que Picard et Weiss [82] se sont intéressés à ce problème dans le cadre de la localisation de sources, pour des outliers d'une cardinalité donnée. Ils donnent une borne fine linéaire pour le nombre d'outliers en fonction du nombre d'observations. Cette borne dépend de la norme L^2 de vecteurs singuliers d'une matrice construite à partir de la matrice d'observation. La réalisation de ce calcul dépend fortement de la structure de son problème et paraît mal adaptée au problème de super-résolution. Kuppinger et coll. [61] ainsi que Stu-

der et coll. [96] étudient le cas où un signal est parcimonieux dans un espace et est corrompu par un bruit parcimonieux dans un espace donné (cas que l'on étudie précisément dans le chapitre suivant). Ce signal est observé par un nombre sous-déterminé d'observations. Ils donnent des conditions basées sur la cohérence mutuelle entre la matrice d'observation du signal parcimonieux et la transformée parcimonieuse du bruit (notion que l'on retrouvera) pour déduire le nombre d'outliers qu'il est possible de pardonner. La dépendance de ce nombre en fonction de la quantité de données n'y est pas plus étudiée que dans [24]. Dans ce chapitre, on montre en utilisant l'équivalence entre reconstruction parcimonieuse et résistance aux outliers qu'une interpolation SR parfaite par minimisation L^1 peut être possible lorsque les données sont contaminées par des outliers. On utilise d'abord une condition équivalente à la garantie de reconstruction parfaite : la propriété de non-concentration, appelée *Null Space Property* (NSP) dans le cadre de la reconstruction parcimonieuse [37, 29]. On détermine aussi une condition de reconstruction parfaite en utilisant la RIP. Enfin, on montre l'importance de considérer la structure des outliers et notamment leur support relativement au support du signal.

Des recherches simultanées sur ce sujet ont été effectuées. [71] fait des développements théoriques similaires et s'intéresse à la robustesse avec des observations aléatoires. Il donne une condition dépendant d'une notion d'angle entre l'espace des outliers et celui des inconnues, ce qui revient à étudier la cohérence mutuelle et donc se baser sur la *Restricted Isometry Property*. [83] étend les résultats de [96] au cas de supports de parcimonie aléatoires. [95] étend les résultats de [96] pour le cas où un bruit d'énergie finie contamine en plus les données.

6.1.2 Définitions

On définit un bruit de type outlier uniquement par le support des données qu'il contamine et on ne fait aucune hypothèse sur les valeurs que ce bruit peut prendre.

Les images BR sont toujours générées par l'application linéaire A du Chapitre 3. Dans le cas d'un bruit de type outlier, on écrit la minimisation L^1 :

$$\operatorname{argmin}_u \|Au - w\|_1 \quad (6.1)$$

avec $w = Au_0 + n_0$. On cherche des conditions sur A pour s'assurer que u_0 est l'unique solution de (6.1) lorsque n_0 est de la forme : $n_0 = n.T$ avec T un vecteur de 0 et 1 représentant un support (le symbole \cdot représente le produit coordonnée à coordonnée).

Dans cette section, on suppose que A est une matrice quelconque de taille $N \times M$ de rang plein (où N et M sont des entiers quelconques tels que $N \geq M$). Les résultats sont ainsi généraux pour les problèmes inverses sur-déterminés. On définit la notion de matrice pardonnante :

Définition 6.1.1 (Matrice pardonnante). *Soit \mathcal{T} un ensemble de supports dans \mathbb{R}^N ($\mathcal{T} \subset \{0, 1\}^N$). A est \mathcal{T} -pardonnante si pour tout $T \in \mathcal{T}$, $n \in \mathbb{R}^N$, $u_0 \in \mathbb{R}^M$, on a :*

$$u_0 = \operatorname{argmin}_u \|Au - (Au_0 + n.T)\|_1 \quad (6.2)$$

et u_0 est l'unique minimiseur.

On répond donc à la question que l'on s'est posée pour une classe d'outliers à support dans \mathcal{T} si l'on arrive à déterminer les conditions (principalement le nombre d'images BR

dans le cas de la super-résolution) pour lesquelles A est \mathcal{T} -pardonnante. En effet, dans le cas de la super-résolution, si A est \mathcal{T} -pardonnante, alors la super-résolution est parfaitement informée pour les outliers à support dans \mathcal{T} .

Lorsque le support des outliers est de cardinal K , il est montré dans [24] que le pardon est équivalent à la capacité de reconstruction parcimonieuse de matrices B dont le noyau est exactement l'image de A ($BA = 0$). On définit les matrices :

Définition 6.1.2 (Matrice Annihilatrice). *Soit A une matrice $N \times M$. On dit que B est un annihilateur de A ou annihile A si B est de taille $(N - M) \times N$, de rang plein et $BA = 0$.*

Le résultat d'équivalence entre pardon et propriété de non concentration s'étend naturellement à des ensembles de supports quelconques \mathcal{T} , ce que l'on va montrer dans le théorème 6.1.2. Pour cela, on étend la définition de ces concepts à des ensembles de supports quelconques.

Définition 6.1.3 (Propriété de Reconstruction Parcimonieuse). *Soit \mathcal{T} un ensemble de supports de \mathbb{R}^N . La matrice B a la \mathcal{T} -Propriété de Reconstruction Parcimonieuse si pour tout $T \in \mathcal{T}$, $x_0 \in \mathbb{R}^N$, on a :*

$$(x_0.T) = \operatorname{argmin}_x \|x\|_1 \text{ t.q. } Bx = B(x_0.T) \quad (6.3)$$

et $x_0.T$ est l'unique minimiseur.

La remarque de [24] peut donc être formulée.

Théorème 6.1.1. *Le \mathcal{T} -pardon de A est équivalent à la \mathcal{T} -propriété de reconstruction parcimonieuse des annihilateurs de A .*

Démonstration. Voir [24] ou utiliser le théorème 6.1.2. □

La super-résolution L^1 est donc robuste aux outliers de support dans \mathcal{T} si et seulement si $\operatorname{Im} A$ vérifie la propriété suivante :

Définition 6.1.4 (Propriété de non concentration (PNC)). *Soient \mathcal{T} un ensemble de supports de \mathbb{R}^N et V un sous espace de \mathbb{R}^N . V a la \mathcal{T} -propriété de non concentration si pour tout $v \in V \setminus \{0\}$ et $T \in \mathcal{T}$:*

$$\|v.T\|_1 < \|v.T^c\|_1 \quad (6.4)$$

où T^c est le complément du support T .

Cette propriété permet de faire des hypothèses différentes sur le support du bruit en considérant des ensembles \mathcal{T} particuliers.

On dit qu'une matrice a la \mathcal{T} -Null Space Property (\mathcal{T} -NSP) si son noyau a la \mathcal{T} -PNC.

On remarque, qu'étant en dimension finie, la PNC implique l'existence d'une constante $\gamma < 1$ telle que pour tout $v \in V$ et tout $T \in \mathcal{T}$:

$$\|v.T\|_1 \leq \gamma \|v.T^c\|_1. \quad (6.5)$$

Cette constante est appelée constante de NSP dans le domaine de la reconstruction parcimonieuse.

6.1.3 Démonstration directe de la caractérisation par la PNC

La littérature nous fournit les équivalences :

$$\text{Pardon} \Leftrightarrow \text{Reconstruction Parcimonieuse}$$

$$\text{Non Concentration} \Leftrightarrow \text{Reconstruction Parcimonieuse}$$

On montre ici directement la troisième équivalence, avec des supports quelconques :

$$\text{Non Concentration} \Leftrightarrow \text{Pardon}$$

Théorème 6.1.2. *Les propositions suivantes sont équivalentes :*

1. A est \mathcal{T} -pardonnante
2. $\text{Im}A$ a la \mathcal{T} -Propriété de Non Concentration.

Démonstration. $1 \Rightarrow 2$: Soit A une matrice \mathcal{T} -pardonnante, et $T \in \mathcal{T}$. Soit $w \in \text{Im}A \setminus \{0\}$, il existe un u_0 tel que $w = Au_0 \neq 0$.

Comme A est pardonnante, u_0 minimise la norme L^1 de l'équation (6.2). On a donc l'inégalité

$$\|n.T\|_1 < \|Au - (w + n.T)\|_1, \quad (6.6)$$

pour tout $n \in \mathbb{R}^N$ et pour tout $u \neq u_0$ sous-optimal. L'inégalité stricte est une conséquence de l'unicité (supposée par hypothèse). En particulier, pour $n = w$ et $u = 2u_0$ ($u \neq u_0$ car $Au_0 \neq 0$), $Au = 2w$ et :

$$\|w.T\|_1 < \|w - w.T\|_1 = \|w.T^c\|_1. \quad (6.7)$$

Cela montre que $\text{Im}A$ satisfait la PNC sur \mathcal{T} .

$2 \Rightarrow 1$: Par hypothèse, $\text{Im}A$ a la PNC sur \mathcal{T} . Soit $u_0 \in \mathbb{R}^M$, $n \in \mathbb{R}^N$ et $T \in \mathcal{T}$. On montre que u_0 est un minimiseur de l'expression de l'équation (6.2). Soit $u \neq u_0$. La norme L^1 est la somme des normes L^1 calculées sur des supports complémentaires :

$$\begin{aligned} f(u) &= \|Au - (Au_0 + n.T)\|_1 \\ &= \|(Au - (Au_0 + n.T)).T\|_1 + \|(Au - (Au_0 + n.T)).T^c\|_1 \\ &= \|(A(u - u_0)).T - n.T\|_1 + \|(A(u - u_0)).T^c\|_1. \end{aligned} \quad (6.8)$$

On utilise l'inégalité triangulaire suivie de la PNC :

$$\begin{aligned} f(u) &\geq \|n.T\|_1 - \|(A(u - u_0)).T\|_1 + \|(A(u - u_0)).T^c\|_1 \\ f(u) &> \|n.T\|_1 = f(u_0). \end{aligned} \quad (6.9)$$

Cette inégalité stricte montre que u_0 est le minimiseur unique de f et que A est \mathcal{T} -pardonnante. \square

Remarque 6.1.1. *On peut interpréter cette équivalence en terme de quantité de bruit. En effet, ce théorème dit que si pour tout bruit supporté par T , sa mesure par la norme L^1 (son « énergie L^1 ») est inférieure à l'énergie des observations non-contaminées, alors le problème est robuste.*

Avec ce résultat, la PNC peut être vérifiée sur des ensembles de supports quelconques et non d'un cardinal fixe, comme cela est souvent fait dans le domaine de la reconstruction parcimonieuse. Par exemple, pour la super-résolution, il peut-être intéressant de considérer les supports contaminant un nombre fixe d'images. Cette hypothèse modélise des situations réelles comme l'apparition d'objets dans la scène, de reflets... Cette généralisation permet d'introduire une notion de structure dans la définition des outliers et de leurs supports, ce qui va jouer un rôle déterminant dans la suite.

Remarque 6.1.2. Dans le contexte de la robustesse aux outliers, on pourrait appeler la PNC, « Propriété de l'Espace Image » (« Image Space Property ») de A pour traduire le concept de Null Space Property de la théorie de la reconstruction parcimonieuse.

6.1.4 Stabilité

Les images BR, comme on l'a vu dans les chapitres précédents, sont contaminées par un bruit blanc aléatoire. Il faut donc s'assurer que la minimisation L^1 est stable par rapport à un tel bruit. Cette stabilité est démontrée dans le cas de la reconstruction parcimonieuse. On donne ici une démonstration similaire directe pour le problème de pardon :

Théorème 6.1.3. Supposons que A est \mathcal{T} -pardonnante. Soit $T \in \mathcal{T}$. Si $w = Au_0 + n.T + n_2$ alors le résultat \tilde{u} de la minimisation (6.1) respecte l'inégalité suivante :

$$\|\tilde{u} - u_0\|_1 \leq 2c \frac{1+\gamma}{1-\gamma} \|n_2.T^c\|_1 \quad (6.10)$$

avec $\gamma < 1$ la constante de PNC de $\text{Im}A$, et c est une constante qui dépend de A .

Démonstration. D'après le théorème 6.1.2, γ existe. En faisant un changement de variable, il suffit de prouver le résultat pour $u_0 = 0$. On peut supposer que n_2 est supporté sur T^c car la partie supportée sur T peut être considérée comme outlier. Comme \tilde{u} est le minimiseur de l'attache aux données L^1 , on a

$$\|A\tilde{u} - n.T - n_2.T^c\|_1 \leq \|Av + n.T + n_2.T^c\|_1, \quad (6.11)$$

pour tout v . On sépare le calcul de la norme de gauche sur des supports complémentaires :

$$\|A\tilde{u} - n.T - n_2.T^c\|_1 = \|(A\tilde{u}).T - n.T\|_1 + \|(A\tilde{u}).T^c - n_2.T^c\|_1. \quad (6.12)$$

On utilise ensuite l'inégalité triangulaire et on choisit $v = 0$:

$$\|n.T\|_1 - \|(A\tilde{u}).T\|_1 + \|(A\tilde{u}).T^c\|_1 - \|n_2.T^c\|_1 \leq \|n.T\|_1 + \|n_2.T^c\|_1, \quad (6.13)$$

ce qui donne finalement.

$$-\|(A\tilde{u}).T\|_1 + \|(A\tilde{u}).T^c\|_1 \leq 2\|n_2.T^c\|_1 \quad (6.14)$$

D'après la PNC de A , on a

$$(1 - \gamma)\|(A\tilde{u}).T^c\|_1 \leq 2\|n_2.T^c\|_1. \quad (6.15)$$

D'après la PNC, on a aussi $\|(A\tilde{u}).T^c\|_1 \geq 1/(1 + \gamma)\|A\tilde{u}\|_1$. L'opérateur A est injectif. Il existe donc une constante c qui dépend de A telle que

$$\|\tilde{u}\|_1 \leq 2c \frac{1 + \gamma}{1 - \gamma} \|n_2.T^c\|_1. \quad (6.16)$$

□

Remarque 6.1.3. L'équation (6.10) implique la stabilité pour toutes les normes

$$\|\tilde{u} - u_0\|_p \leq C \|n_2\|_q. \quad (6.17)$$

où C est une contante qui dépend de p, q, A et \mathcal{T} .

Ce résultat permet dans le cas de la super-résolution de déduire qu'elle est bien informée dans le cas d'une contamination par des outliers et par du bruit gaussien si l'opérateur est pardonnant.

6.2 Conditions de reconstruction parfaite par SR en présence d'outliers

Dans cette section, on cherche des conditions suffisantes de pardon pour l'opérateur de SR. La matrice A est donc de taille $Nl^2 \times (Ml)^2$. On commence par une étude naïve utilisant un nombre d'outliers donné. On remarque ensuite que les bornes sont améliorées si l'on considère une structure particulière sur les outliers : une mesure de parcimonie utile est le nombre d'images contaminées, et non le nombre d'échantillons.

6.2.1 Conditions suffisantes pour le K -pardon

Dans cette section, \mathcal{T} est l'ensemble des supports de cardinal K . La propriété de non-concentration permet de trouver des conditions suffisantes de reconstruction sur le nombre d'images BR lorsque K outliers contaminent les données. Ces conditions font intervenir le conditionnement de A .

Définition 6.2.1 (Conditionnement L^p). *Le conditionnement L^p d'un opérateur A est définit par :*

$$\kappa_{A,p} = \frac{\sup_{\|u\|_p=1} \|Au\|_p}{\inf_{\|u\|_p=1} \|Au\|_p} \quad (6.18)$$

En utilisant la PNC, on montre la propriété suivante.

Proposition 6.2.1. *Soit A une matrice de super-résolution. Soit $\kappa_{A_{\mathcal{T}^c},1}$, le conditionnement L^1 maximum des $A_{\mathcal{T}^c}$ (les A dont les lignes sont restreintes à \mathcal{T}^c). Supposons :*

$$N > K(M^2 \kappa_{A_{\mathcal{T}^c},1} + \frac{1}{l^2}) \quad (6.19)$$

Alors A est K -pardonnante.

Démonstration. Soit T un support de cardinal K . On cherche une condition suffisante pour que

$$\frac{\|A_T|u\|_1}{\|A_{T^c}|u\|_1} < 1 \quad (6.20)$$

soit vrai pour tous les supports T de taille K . On commence par borner la norme d'opérateur L^1 de $A_T|$ (A restreint à T). Soient a_i les lignes de A :

$$\frac{\|A_T|u\|_1}{\|u\|_1} = \frac{\sum_{i \in T} | \langle a_i, u \rangle |}{\|u\|_1} \leq \frac{\sum_{i \in T} \sum_j |a_{i,j} u_j|}{\|u\|_1}. \quad (6.21)$$

On a $|a_{i,j}| \leq 1$ car tous les coefficients de A sont des échantillons d'un sinus cardinal. En conséquence :

$$\begin{aligned} \frac{\|A_T|u\|_1}{\|u\|_1} &\leq \frac{\sum_{i \in T} \sum_j |u_j|}{\|u\|_1} \\ &\leq K. \end{aligned} \quad (6.22)$$

On borne ensuite le rapport $\frac{\|A_T|u\|_1}{\|A_{T^c}|u\|_1}$. On utilise le conditionnement L^1 de $A_{T^c}|$, noté $\kappa_{A_{T^c}|,1}$. On peut ainsi écrire :

$$\begin{aligned} \frac{\|A_T|u\|_1}{\|A_{T^c}|u\|_1} &\leq \frac{K\|u\|_1}{\|A_{T^c}|u\|_1} \\ &\leq K \left(\inf \frac{\|A_{T^c}|u\|_1}{\|u\|_1} \right)^{-1} \\ &\leq K \frac{\kappa_{A_{T^c}|,1}}{\|A_{T^c}\|_{op,1}}. \end{aligned} \quad (6.23)$$

On utilise le fait que la norme d'opérateur L^1 , $\|A_{T^c}\|_{op,1}$ peut être bornée inférieurement par des valeurs particulières. On remarque que l'opérateur SR transforme les images HR constantes en images BR constantes ayant la même intensité. En conséquence, $\|A_{T^c}\|_{op,1} \geq (Nl^2 - K)/(Ml)^2$ et :

$$\frac{\|A_T|u\|_1}{\|A_{T^c}|u\|_1} \leq K(Ml)^2 \frac{\kappa_{A_{T^c}|,1}}{Nl^2 - K}. \quad (6.24)$$

On en déduit la proposition. \square

Le conditionnement L^1 est une quantité difficile à calculer en général, car il faut borner inférieurement la valeur absolue d'une somme. Pour certains problèmes, il est possible de le faire. Par exemple, on peut borner inférieurement, la norme L^1 d'une somme d'exponentielles complexes [59].

Grâce à l'équivalence entre propriété de reconstruction parcimonieuse et pardon, on peut utiliser la *Restricted Isometry Property* (RIP, [24]) pour trouver une condition suffisante de K -pardon en utilisant des normes L^2 . On rappelle la définition de la RIP.

Définition 6.2.2 (Restricted Isometry Property). *B a la Restricted Isometry Property d'ordre J avec constante $\delta \in]0, 1[$ si pour tout $x \in \mathbb{R}^{N(l \times l)}$ et pour tout support T tel que $|T| = J$*

$$(1 - \delta)\|x.T\|_2 \leq \|B(x.T)\|_2 \leq (1 + \delta)\|x.T\|_2. \quad (6.25)$$

Montrer la RIP d'ordre $J = K + K'$ avec constante $\delta < \frac{\sqrt{K'} - \sqrt{K}}{\sqrt{K'} + \sqrt{K}}$ pour B donne la capacité de reconstruction parcimonieuse à B pour les supports de taille K (voir [29, 31]). Si de plus B annihile A , A est K -pardonnante. On peut facilement construire des annihilateurs de A , on peut donner une version de la RIP qui ne dépend que de A :

Proposition 6.2.2. *Si pour tout T de cardinal J :*

$$\frac{\|A(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} \leq \sqrt{\delta} \quad (6.26)$$

alors il existe B ayant la RIP d'ordre J et constante δ telle que $\ker B = \text{Im} A$.

Démonstration. Étant donné une matrice A , on définit B comme la projection orthogonale sur $(\text{Im} A)^\perp$: $B = P_{(\text{Im} A)^\perp} = I - A(A^H A)^{-1} A^H$. En restreignant B à son espace image (qui est de dimension $(N - M^2)l^2$), on construirait une matrice de rang plein et donc annihilatrice de A .

On réexprime l'équation (6.25) en fonction de A . On commence par l'élever au carré :

$$(1 - \delta)^2 \|x.T\|_2^2 \leq \|B(x.T)\|_2^2 \leq (1 + \delta)^2 \|x.T\|_2^2. \quad (6.27)$$

D'après le théorème de Pythagore,

$$\|x.T\|_2^2 = \|A(A^H A)^{-1} A^H(x.T)\|_2^2 + \|B(x.T)\|_2^2 \quad (6.28)$$

On obtient en remplaçant dans l'équation (6.27) :

$$(1 - \delta)^2 \leq 1 - \frac{\|A(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} \leq (1 + \delta)^2 \quad (6.29)$$

On en déduit : si pour tout T de cardinal J :

$$\frac{\|A(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} \leq \sqrt{\delta} \quad (6.30)$$

alors B a la RIP d'ordre J et constante δ . □

Proposition 6.2.3. *Supposons*

$$N > M^2 C_1^{-1} K \kappa_{A,2}^4 \quad (6.31)$$

où $C_1 = 0.0670$. Alors A est K -pardonnante.

Démonstration. Soit un support T de cardinal J . On montre d'abord que $\|A_T^H\|_{op,2} = \|A_T\|_{op,2} \leq \sqrt{J}$. En utilisant l'inégalité de Cauchy-Schwarz :

$$\frac{\|A_T u\|_2^2}{\|u\|_2^2} = \frac{\sum_{i \in T} |\langle a_i, u \rangle|^2}{\|u\|_2^2} \leq \frac{\sum_{i \in T} \|a_i\|_2^2 \|u\|_2^2}{\|u\|_2^2}. \quad (6.32)$$

On a $\|a_i\|_2 \leq 1$ car ce sont des sinus cardinaux. En conséquence :

$$\frac{\|A_T u\|_2}{\|u\|_2} \leq \sqrt{J} \quad (6.33)$$

On borne maintenant le ratio :

$$\begin{aligned}
 \frac{\|A(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} &\leq \sigma_{max} \frac{\|(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} \\
 &\leq \sigma_{max} \sigma_{min}^{-2} \|A_T^H\|_2 \\
 &\leq \frac{\kappa_{A,2}^2 \sqrt{J}}{\sigma_{max}}
 \end{aligned} \tag{6.34}$$

où σ_{min} et σ_{max} sont les valeurs singulières minimale et maximale de A . En remplaçant avec une valeur admissible de δ , on trouve la condition :

$$\frac{\kappa_{A,2}^4 (K + K')}{\sigma_{max}^2} \leq \frac{\sqrt{K'} - \sqrt{K}}{\sqrt{K} + \sqrt{K'}}. \tag{6.35}$$

On prend $K' = 3K$ (que l'on a trouvé optimal pour la constante C_1). On obtient alors la condition :

$$\frac{\kappa_{A,2}^4}{\sigma_{max}^2} \leq \frac{C_1}{\sqrt{K}} \tag{6.36}$$

$\sigma_{max} \geq \frac{\|Au\|_2}{\|u\|_2}$ car σ_{max} est la norme d'opérateur de A . Si on prend pour u une image constante, on obtient : $\sigma_{max} \geq \sqrt{N/M^2}$. Finalement, la condition devient,

$$N > M^2 C_1^{-1} K \kappa_{A,2}^4 \tag{6.37}$$

□

Cette inégalité est utile car elle dépend du conditionnement **global** du problème de SR et la constante est facilement calculable (même si elle est grande). De plus, on sait comment se comporte le conditionnement global. Dans le chapitre 3, il est montré que le conditionnement $\kappa_{A,2}$ tend vers 1 pour un grand nombre d'images avec un mouvement aléatoire. On peut de plus montrer que l'on ne peut pas avoir de borne meilleure que linéaire en fonction du nombre d'outliers en exhibant le cas particulier d'outliers contaminant le même pixel (on le montrera pour le cas particulier du débruitage multi-image en utilisant la PNC dans le chapitre 7). On le montre ici avec un contre exemple pour un cas particulier de super-résolution.

Proposition 6.2.4. *Pour les problèmes de super-résolution 1D, avec un facteur $M = 2$ et un nombre d'observations $N = 2P > 2$, on ne peut pas avoir K -pardon en général si $K > N/4$.*

Démonstration. On construit un contre-exemple. Supposons que les translations des signaux BR sont 0, 1, ..., 0, 1 respectivement. On a donc P observations avec une translation 0 et P avec une translation 1. Les signaux BR coïncident exactement avec la grille HR. Dans ce cas, la reconstruction pour un échantillon HR est la médiane des P observations disponibles pour cet échantillon. Pour que cette médiane donne le bon résultat, il faut qu'au moins $P/2$ observations par pixel ne comporte pas d'outliers. En effet, dans le pire des cas, tous les outliers prennent une même valeur différente du signal. □

On peut étendre ces bornes inférieures au cas 2D :

Proposition 6.2.5. *Pour les problèmes de super-résolution 2D, avec un facteur $M = 2$ et un nombre d'images $N = M^2 P$, on ne peut pas avoir K -pardon en général si $K > \frac{N}{2M^2}$.*

Démonstration. On utilise exactement le même argument que pour la proposition 6.2.4 précédente. \square

Avec ces développements, on trouve donc asymptotiquement la borne linéaire qui est optimale. La borne asymptotique de l'équation (6.37) est de $N/15M^2$ outliers à comparer à la limite $N/2M^2$. Cependant, on voit que ces bornes ne sont pas bonnes quantitativement, le nombre d'images nécessaires pour remplir la condition augmente avec le nombre d'outliers, mais on se rend compte que la borne qui utilise le contre-exemple que l'on vient d'exhiber ne change pas si l'on considère que les outliers contaminent toute une image. En utilisant cette remarque, on améliore les bornes.

6.2.2 Étude d'une structure d'outliers particulière

Ici, on illustre les bénéfices de pouvoir faire une hypothèse sur la structure des outliers. Soit \mathcal{T} l'ensemble des supports qui contaminent au plus N_c images BR. Comme précédemment, on essaie de trouver une condition suffisante pour la PNC avec $T \in \mathcal{T}$. Cette structure permet d'avoir $K = N_c l^2$ outliers tant qu'ils ne contaminent pas plus de N_c images. Soit \mathcal{S} l'ensemble des indices des images BR contaminées ($|\mathcal{S}| = N_c$). On montre la proposition suivante :

Proposition 6.2.6. *Une condition suffisante de \mathcal{T} -pardon est*

$$N_c C_2 C_3 \|(A_{T^c}^H A_T)^{-1}\|_{op,1} < 1 \quad (6.38)$$

avec C_2, C_3 des constantes qui dépendent de la taille des images.

Démonstration. On commence par borner plus strictement la norme d'opérateur.

$$\begin{aligned} \frac{\|A_T u\|_1}{\|u\|_1} &\leq \sum_{i \in \mathcal{S}} \|A_i\|_{op,1} \\ &\leq \sum_{i \in \mathcal{S}} \|S Q_i\|_{op,1} \\ &\leq C_2 N_c \end{aligned} \quad (6.39)$$

où C_2 est une borne supérieure pour $\|A_i\|_{op,1}$. C_2 est la norme L^1 maximale qu'une colonne de A_i peut atteindre. En utilisant cette borne, on a :

$$\begin{aligned} \frac{\|A_T u\|_1}{\|A_{T^c} u\|_1} &\leq \frac{\|A_T u\|_1 \|u\|_1}{\|u\|_1 \|A_{T^c} u\|_1} \\ &\leq \frac{\|u\|_1 C_2 N_c}{\|A_{T^c} u\|_1} . \end{aligned} \quad (6.40)$$

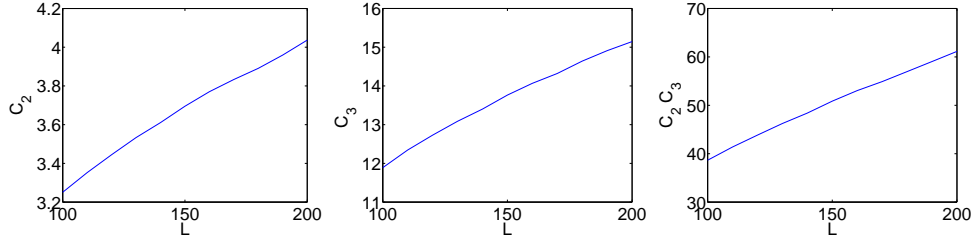


FIGURE 6.1. Constantes pour la SR translationnelle. (a) Évaluation de C_2 en fonction de l , (b) évaluation de C_3 en fonction de l , (c) évaluation de C_2C_3 en fonction de l .

On introduit la pseudo-inverse $A_{T^c|}^\dagger = (A_{T^c|}^H A_{T^c|})^{-1} A_{T^c|}^H$ (on rappelle que $A_{T^c|}$ est de rang plein presque sûrement si $N - N_c > M^2$) :

$$\begin{aligned} \sup_u \frac{\|u\|_1}{\|A_{T^c|} u\|_1} &= \sup_{v \in \text{Im} A_{T^c|}} \frac{\|A_{T^c|}^\dagger v\|_1}{\|v\|_1} \\ &\leq \|(A_{T^c|}^H A_{T^c|})^{-1}\|_{op,1} \sup_{v \in \text{Im} A_{T^c|}} \frac{\|A_{T^c|}^H v\|_1}{\|v\|_1} \\ &\leq \|(A_{T^c|}^H A_{T^c|})^{-1}\|_{op,1} C_3 \end{aligned} \quad (6.41)$$

où C_3 est le maximum de la norme L^1 des colonnes de $Q_i^H S^H$ (ou des lignes des A_i). Pour des signaux 1D, la norme L^1 des *sinc* est grossièrement bornée par le logarithme de la taille de son support. En combinant, les équations (6.39) et (6.41), on obtient le résultat. \square

On évalue expérimentalement figure 6.1, les constantes C_2, C_3 et le produit C_2C_3 . Pour cela, pour chaque taille d'image, on a généré aléatoirement 500 colonnes des matrices concernées. On en a calculé la norme L^1 et gardé la valeur maximale sur toutes les expériences.

Le terme $\|(A_{T^c|}^H A_{T^c|})^{-1}\|_1$ ne peut pas être borné sans hypothèse sur les mouvements. On sait qu'en moyenne $(A_{T^c|}^H A_{T^c|})^{-1} \sim \frac{1}{N} I$ (chapitre 3) lorsque le support T est fixé et $N \rightarrow \infty$. De plus, on peut estimer numériquement les constantes. Asymptotiquement, la contrainte devient $N_c < C_4 N$ (pour $l = 200$, on a $C_4 = 60$). Ce qui est bien mieux que le résultat précédent (l'équation (6.31)) où la constante aurait valu 97000 pour $l = 200$. La constante C_4 reste grande pour des cas pratiques, mais on sait qu'elle ne peut être inférieure à 8. On en a donc un encadrement assez précis.

6.2.3 Algorithmes et exemples

On a utilisé un algorithme de moindres carrés itératif pondéré pour minimiser la norme L^1 [31]. Il est montré dans [31] que cet algorithme converge vers la solution K -parcimonieuse si A a la $(K + \frac{2\gamma}{1-\gamma})$ -propriété de reconstruction parcimonieuse (avec γ la constante de PNC et lorsque les poids sont choisis de manière adéquate (avec une régularisation des poids $\epsilon_n \rightarrow 0$)). On utilise cet algorithme en utilisant l'algorithme de SR L^2 présenté au cha-

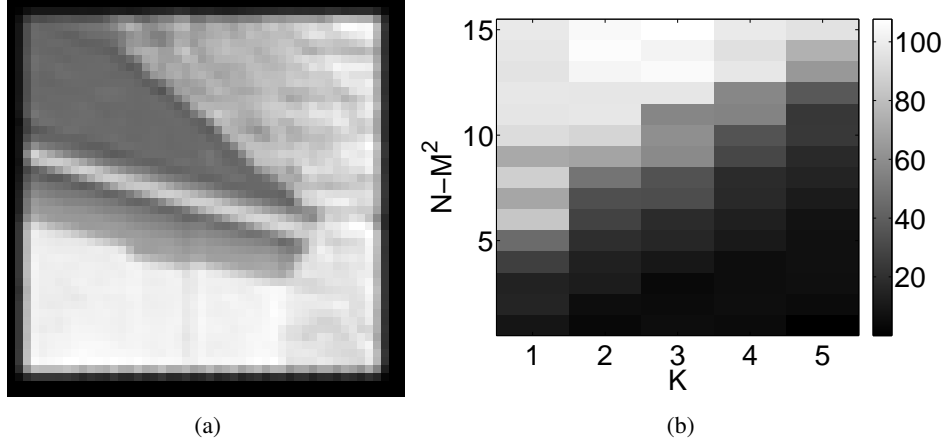


FIGURE 6.2. Robustesse aux outliers. (a) Image HR utilisée pour toutes les expériences, (b) 10% centile du PSNR (en dB) en fonction du nombre d'images contaminées K et du nombre d'images $N - M^2$.

pitre 3. On construit les itérations équivalentes à celles de [31] pour notre problème :

$$\begin{aligned}
 u_{n+1} &= \operatorname{argmin}_u \|\Omega_n(Au - w)\|_2^2 \\
 z_{n+1} &= Au_{n+1} - w \\
 r_{n+1} &= \text{decreasing sort of } \operatorname{abs}(z_{n+1}) \\
 \epsilon_{n+1} &= \min(\epsilon_n, r_{n+1}(K + 1)) \\
 \Omega_{n+1} &= \operatorname{diag} \left([z_{n+1}^2 + \epsilon_{n+1}^2]^{-1/4} \right)
 \end{aligned} \tag{6.42}$$

On a choisi cet algorithme car il converge rapidement (quelques itérations en pratique) et la convergence peut être vérifiée en étudiant les variations de ϵ .

La figure 6.2 montre une évaluation expérimentale du nombre d'images nécessaires pour une reconstruction parfaite lorsque K images sont contaminées. Pour chaque K, N , on calcule le PSNR pour le résultat de 30 expériences avec différents mouvements. On montre le 10-ème centile (90% des reconstructions ont un meilleur PSNR). Dans la figure 6.3, on insère une image BR composée de la valeur absolue d'un bruit gaussien de variance 125 (les pixels ont des valeurs dans $[0, 255]$). Dans ce cas, 6 images non bruitées de plus donnent une reconstruction parfaite de l'image de haute résolution. Dans la figure 6.4, même avec plus d'images contaminées (4 images BR bruitées sur 8 images), si la localisation spatiale des outliers est différente entre les images BR, la minimisation L^1 est robuste.

6.3 Conclusion

On a étudié les capacités de rejet d'outliers de la super-résolution L^1 de manière théorique. On a montré que si le ratio images non bruitées / images bruitées est suffisamment grand, on peut garantir la reconstruction parfaite. L'interpolation de super-résolution est donc parfaitement informée dans ces cas là. Le résultat de stabilité garantit qu'elle reste

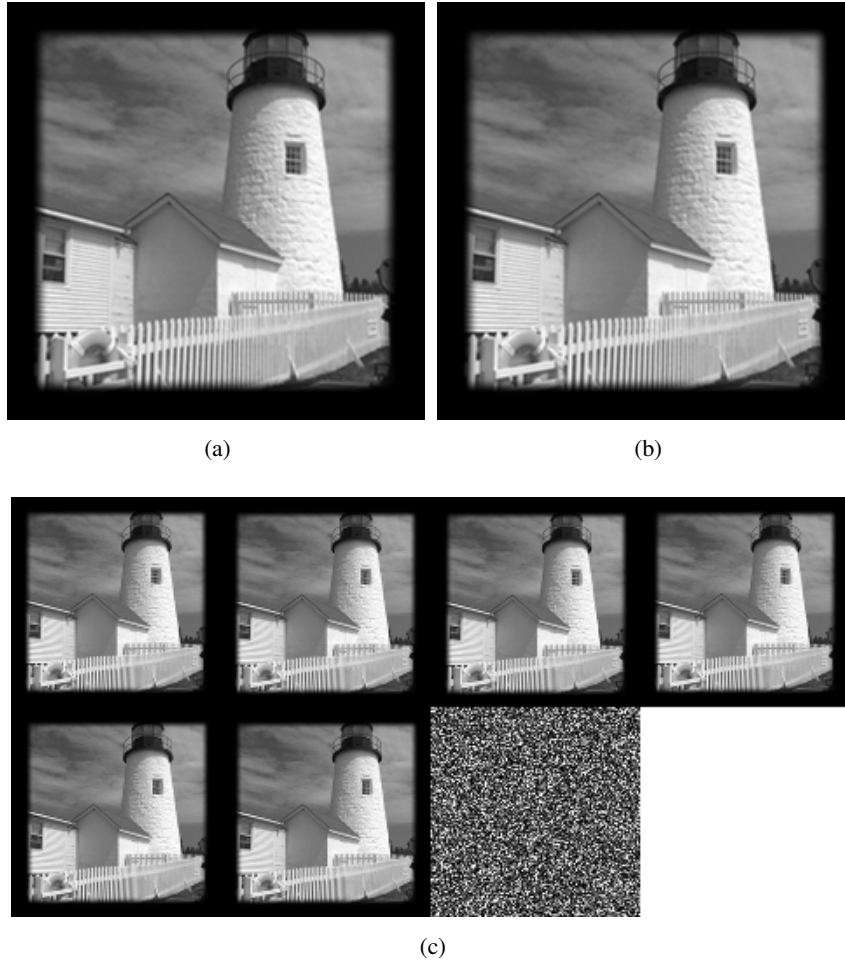


FIGURE 6.3. Interpolation SR L^1 pour $M = 2$ et $N = 7$ avec une seule image contaminée. (a) Image HR, (b) image reconstruite, (c) images BR (outliers sur la dernière image).

bien informée dans le cas d'un bruit d'énergie finie additionnel. On a donné des bornes théoriques pour assurer cette reconstruction parfaite. En pratique, moins d'images semblent nécessaires pour l'obtenir. On a de plus remarqué que si les supports des outliers ne se recouvrent pas, la capacité de pardon augmente, on va exploiter cette idée dans le cadre plus général des méthodes $L^1 - L^1$ dans le chapitre suivant. On remarque aussi que, bien que l'on ait utilisé la théorie de la reconstruction parcimonieuse, les résultats dit de *compressed sensing* ne sont pas valables dans le cas de la robustesse de la super-résolution multi-images. En effet, le résultat le plus remarquable de cette théorie est que pour des observations aléatoires, plus on dispose de données, plus le pourcentage d'outliers que l'on peut pardonner est grand (borne meilleure que linéaire). Dans notre cas, ce pourcentage est constant (borne linéaire). En terme de robustesse, si on s'attend à ce qu'une partie de l'image soit contaminée dans un pourcentage constant du nombre d'image, alors ajouter beaucoup d'images n'est pas utile, on pourra juste tenter de vérifier, si la super-résolution sera robuste ou non. Si le nombre d'images contaminées est fixe, ajouter plus d'images, pourra augmenter la

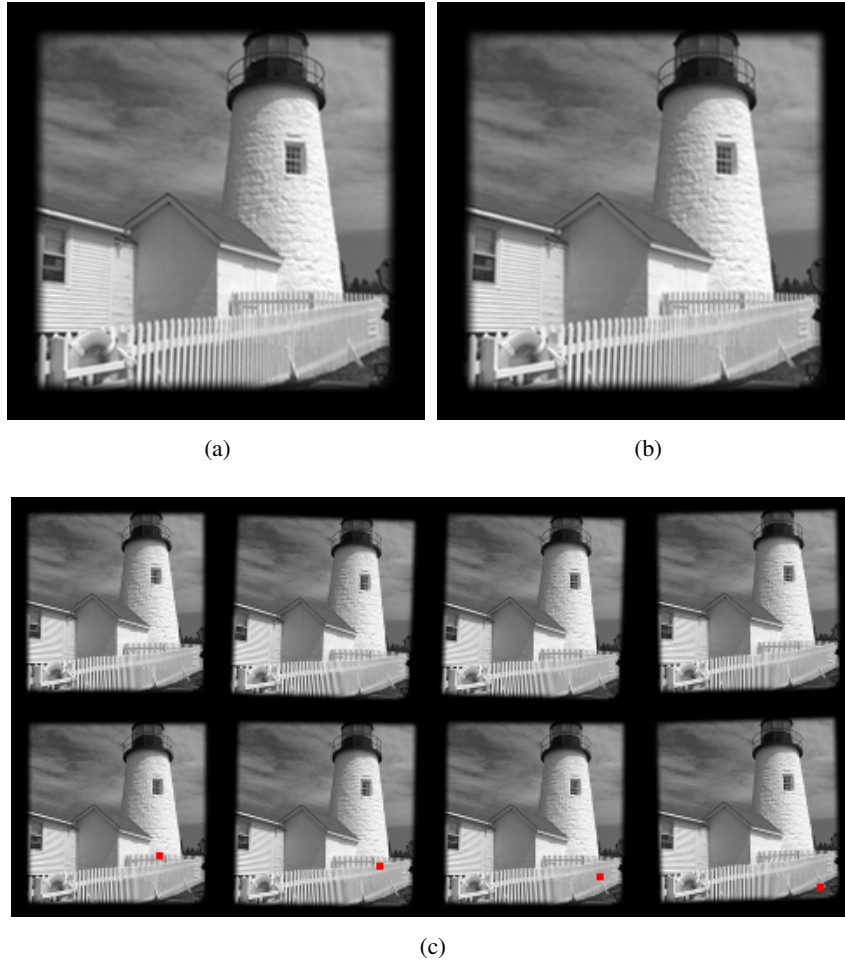


FIGURE 6.4. Interpolation SR L^1 pour $M = 2$ et $N = 8$. (a) Image HR idéale, (b) image reconstruite, (c) images BR (Les outliers simulent des pixels saturés (carrés rouges) sur les 4 dernières).

robustesse de la super-résolution.

CHAPITRE 7

EFFET DE LA RÉGULARISATION PARCIMONIEUSE SUR LA ROBUSTESSE DE LA SUPER-RÉSOLUTION L^1

Dans ce chapitre, on étudie l'effet de l'ajout d'un terme de régularisation au problème du chapitre 6. On commence par inclure ce nouveau terme dans la théorie du pardon. On étudie ensuite le cas des matrices aléatoires afin de faire le lien avec la littérature. Enfin, on s'intéresse au cas du traitement multi-image. On choisit les outils théoriques qui permettent d'expliquer les résultats de robustesse des méthodes L^1 -TV. On trouve que la structure des supports de parcimonie considérés joue un rôle déterminant. Avec des hypothèses appropriées, on arrive à déterminer la valeur du paramètre de régularisation qui permet une reconstruction parfaite.

7.1 Extension du pardon au cas régularisé

L'approximation parcimonieuse est un outil très utilisé pour inverser des systèmes mal conditionnés ou mal posés. Beaucoup d'applications bénéficient de cet outil. De l'acquisition compressée [22] au débruitage des images [38, 114] et à la synthèse de textures [99], la parcimonie d'un signal permet de le reconstruire alors que l'information disponible est plus faible que ce qui est nécessaire a priori. Il a été montré que les images naturelles peuvent avoir des représentations parcimonieuses dans certains dictionnaires [64, 66, 15]. Un dictionnaire est une collection de composantes élémentaires d'une image (ex. ondelettes redondantes).

L'approximation parcimonieuse a principalement été faite sur les signaux. On a montré précédemment que cette même hypothèse sur le bruit permet de déterminer la robustesse aux outliers de la super-résolution L^1 . Connaissant les bénéfices de l'hypothèse de parcimonie sur les signaux, la question de la faire à la fois sur le signal et sur le bruit se pose. Est-ce qu'ajouter une hypothèse de parcimonie sur le signal améliore la robustesse de la super-résolution L^1 ?

Cette question a été étudiée dans [96, 61, 95] avec un modèle d'observation différent : le nombre d'observations du signal parcimonieux est sous-déterminé. Il a été montré qu'un compromis entre la parcimonie du signal et la parcimonie du bruit doit être fait pour garantir la reconstruction parfaite de certains algorithmes (L^0 approché, L^1 et $L^2 - L^1$) et de certaines matrices d'observation. En revanche, la question du paramètre de régularisation n'est pas traitée. De plus, l'effet de ces méthodes sur la super-résolution multi-image n'est pas étudié de manière théorique.

La régularisation parcimonieuse est déjà utilisée pour la super-résolution mono-image [78]. La régularisation par TV [68], populaire pour la super-résolution, peut être interprétée comme la version convexe d'une régularisation de type gradient parcimonieux.

7.1.1 Hypothèses de parcimonie sur les images naturelles

Pour une image u_0 , on dit que u_0 est parcimonieuse en analyse pour une transformée (dite parcimonieuse) particulière Ψ si Ψu_0 est un vecteur parcimonieux. Une large gamme d'hypothèses parcimonieuses (la fonction Ψ , en général, une fonction linéaire) ont été considérées pour les images naturelles. On peut citer les décompositions en ondelettes, les projections dans des dictionnaires d'images (patch(es)) [47, 38]. Une autre supposition fréquente sur les images est le fait qu'elles minimisent la norme L^1 du gradient (la variation totale anisotrope). Interprétée selon la théorie de la reconstruction parcimonieuse, minimiser cette variation totale est équivalent à imposer de la parcimonie sur le gradient. Dans la section sur la super-résolution, nous allons utiliser cette hypothèse comme exemple pratique.

Dans la suite, on étudie la minimisation

$$\operatorname{argmin}_u \|Au - w\|_1 + \lambda \|\Psi u\|_1 \quad (7.1)$$

où Ψ est la transformée parcimonieuse.

7.1.2 Définitions

Soit \mathcal{T} une famille de supports, A une matrice $N \times M$ de rang plein. On étend la définition de pardon au cas régularisé. Soit la transformée parcimonieuse Ψ une matrice de taille $N' \times M$. On veut étudier le cas où A est \mathcal{T} -pardonnante sur les solutions régulières u_0 où Ψu_0 est \mathcal{L} -parcimonieux pour des supports de parcimonie de signal \mathcal{L} . Idéalement, on aimerait étendre le concept comme suit :

Définition 7.1.1 ((L^1, L^0) Matrice pardonnante sous hypothèse parcimonieuse). *Soit \mathcal{T} un ensemble de supports de \mathbb{R}^N et \mathcal{L} un ensemble de supports de $\mathbb{R}^{N'}$. On dit que la matrice A est \mathcal{T} -pardonnante sur les solutions Ψ, \mathcal{L} -parcimonieuses si pour tout $T \in \mathcal{T}, n \in \mathbb{R}^N, u_0 \in \mathbb{R}^M$ tels que $\Psi u_0 = \alpha_0 \in \mathcal{L} \cdot \mathbb{R}^{N'}$, on a :*

$$u_0 = \operatorname{argmin}_u \|Au - (Au_0 + n.T)\|_1 \text{ t.q. } \Psi u \in \mathcal{L} \cdot \mathbb{R}^{N'} \quad (7.2)$$

et u_0 est l'unique minimum.

Cependant, le manque d'algorithme pratique pour résoudre exactement ce problème nous pousse à étudier la relaxation convexe de celui-ci. Imposer l'appartenance du signal à $\mathcal{L} \cdot \mathbb{R}^{N'}$, rend une partie du problème combinatoire. Il a été montré dans [32] que ce problème appartient à la classe des problèmes NP-difficiles, c'est-à-dire que l'on ne connaît pas de moyen de le résoudre en temps polynomial (par rapport à la taille des données). Pour le traitement des images, des algorithmes gloutons (ex. : *Orthogonal Matching Pursuit* [96]) sont utilisés pour essayer de résoudre approximativement de tels problèmes. Les propriétés de convergences de ces algorithmes peuvent alors être étudiées.

On a choisi de garder l'approche et le cadre théorique du chapitre précédent. De cette manière, on étudie seulement les minima de la fonctionnelle de l'équation (7.1) et les résultats ne dépendent pas de la manière de la minimiser. On va donc considérer la régularisation parcimonieuse L^1, L^1 , qui est la version convexe de la définition L^1, L^0 précédente. On présente sur la figure 7.1, un schéma classique qui illustre la différence entre les minimisations L^2, L^1 et L^0 . On voit que, dans ce cas, la minimisation de l'attache aux données L^1

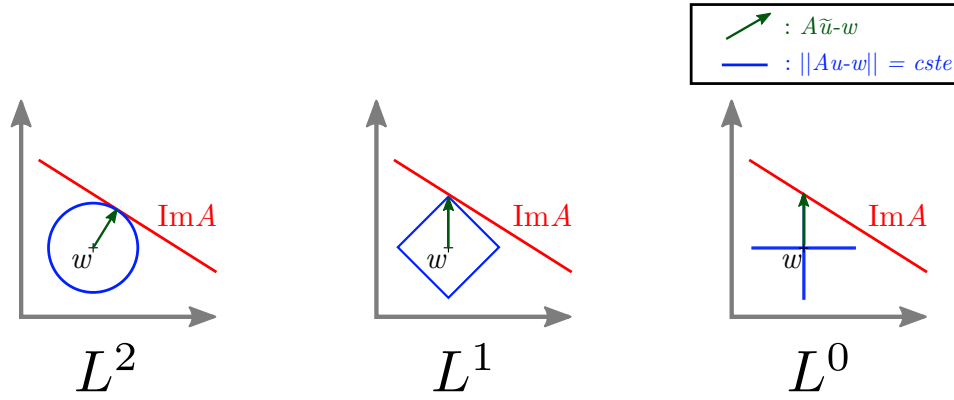


FIGURE 7.1. Minimisations L^2 , L^1 et L^0 . La solution L^1 est identique à la solution L^0

donne la même solution parcimonieuse que la minimisation L^0 pour $A\tilde{u} - w$: la solution a une seule composante non nulle. De plus, on voit que la sphère de norme L^1 constante est l'enveloppe convexe de l'intersection de la sphère de norme L^0 avec la boule unité L^2 . Il est donc légitime de chercher les situations où la version convexe du problème donne le même résultat.

Définition 7.1.2 ((L^1, L^1) Matrice pardonnante sous hypothèse parcimonieuse). Soit $\lambda \in \mathbb{R}$. Soit \mathcal{T} un ensemble de supports de \mathbb{R}^N et \mathcal{L} un ensemble de supports de $\mathbb{R}^{N'}$. On dit que A est \mathcal{T} -pardonnante sur les solutions $\lambda\Psi$, \mathcal{L} -parcimonieuses par optimisation convexe si pour tout $T \in \mathcal{T}$, $n \in \mathbb{R}^N$, $u_0 \in \mathbb{R}^M$ tels que $\Psi u_0 = \alpha_0 \in \mathcal{L}$, on a :

$$u_0 = \operatorname{argmin}_u \|Au - (Au_0 + n.T)\|_1 + \lambda \|\Psi u\|_1 \quad (7.3)$$

et u_0 est l'unique minimiseur.

Remarque 7.1.1. Lorsque l'on effectue la minimisation (7.3), on appelle λ le paramètre de régularisation. Ce paramètre est considéré comme le niveau de régularisation, c'est-à-dire la quantité de régularité que l'on impose sur l'image. Idéalement, on aimerait pouvoir définir la transformée parcimonieuse indépendamment de sa constante de normalisation (λ), comme la capacité de pardon dépend de cette normalisation, on l'a intégré comme paramètre de la définition. Au final, on pourra dire que A est pardonnante sur les solutions parcimonieuses s'il existe un λ qui donne la minimisation (7.3)

Remarque 7.1.2. L'équation (7.3) est équivalente à : u_0 est l'unique minimiseur de

$$\min_u \|A_r^\lambda u - w_r\|_1 \quad (7.4)$$

$$\text{où } A_r^\lambda = \begin{pmatrix} A \\ \lambda\Psi \end{pmatrix} \text{ et } w_r = \begin{pmatrix} Au_0 \\ \lambda\Psi u_0 \end{pmatrix} + \begin{pmatrix} n.T \\ -\lambda\Psi u_0 \end{pmatrix} = \begin{pmatrix} Au_0 + n.T \\ 0 \end{pmatrix}$$

Remarque 7.1.3. D'après la remarque précédente, s'il existe un λ , tel que A_r^λ pardonne les outliers à supports dans $\mathcal{T}' = \left\{ \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix}, \tau_1 \in \mathcal{T}, \tau_2 \in \mathcal{L} \right\}$, on dit alors que A_r^λ est

\mathcal{T}' -pardonnante ou \mathcal{T}, \mathcal{L} -pardonnante). On conclut que la matrice A est \mathcal{T} -pardonnante sur les solutions $\lambda\Psi, \mathcal{L}$ -parcimonieuses.

Remarque 7.1.4. D'après la remarque précédente, si $\text{Im}A_r^\lambda$ a la \mathcal{T}' -PNC, alors A est \mathcal{T} -pardonnante $\lambda\Psi, \mathcal{L}$ -parcimonieuses. On montre dans le théorème 7.1.2 que c'est une équivalence dans le cas où Ψ est inversible.

Quand \mathcal{T} et \mathcal{L} sont les familles de supports de cardinal K et S respectivement. On parle de K, S -pardon et de K, S -PNC.

7.1.3 Propriété de Reconstruction Exacte

Lorsque l'on veut étudier une instance particulière du problème, il est utile d'avoir une condition de reconstruction qui ne dépend que de cette instance du problème. On trouve une telle propriété dans [25] où elle est appelée propriété de reconstruction exacte. On peut aussi la trouver dans la littérature sous l'appellation : existence d'un certificat dual [21, 116]. On peut montrer le théorème suivant en utilisant les propriétés du sous-gradient au minimum de la fonction à minimiser. Elle est utile pour l'étude d'instances particulières du problème. On produit une preuve géométrique directe (calquée directement sur celle de [25]) dans le cas du pardon.

Théorème 7.1.1 (Propriété de Reconstruction Exacte (PRE)). *Soit A, T, n, u_0 une instance du problème (6.1). Une condition suffisante pour que u_0 en soit l'unique minimiseur est l'existence d'un vecteur $v \in \mathbb{R}^N$ tel que :*

1. $v \in \ker A^H$
2. $v.T = \text{sign}(n.T), \|v.T^c\|_\infty < 1$

Démonstration. Parce que l'on peut toujours faire le changement de variable $u' = u - u_0$, il suffit de vérifier que, pour tout u , on a :

$$\|Au - n.T\|_1 \geq \|n.T\|_1 \quad (7.5)$$

En utilisant les propriétés de la norme infinie de v , on a :

$$\|Au - n.T\|_1 \geq \langle v, n.T - Au \rangle \quad (7.6)$$

que l'on décompose :

$$\|Au - n.T\|_1 \geq \langle v, n.T \rangle - \langle v, Au \rangle \quad (7.7)$$

Comme $v.T = \text{sign}(n.T)$, on a $\langle v, n.T \rangle = \|n.T\|_1$ et

$$\|Au - n.T\|_1 \geq \|n.T\|_1 - \langle A^H v, u \rangle \quad (7.8)$$

avec $A^H v = 0$, ce qui mène au résultat. \square

Cette propriété sera utile dans l'analyse fine du pardon pour la SR régularisée, pour les supports de parcimonies structurés dans la section 7.3.

7.1.4 Caractérisation par la propriété de non concentration

On étend le théorème 6.1.2 au cas régularisé, on préserve l'équivalence si Ψ est inversible.

Théorème 7.1.2. *Si $N' = M$ et Ψ est **inversible**, les propositions suivantes sont équivalentes (on rappelle que \mathcal{T}' est l'ensemble des concaténations des supports de \mathcal{T} et \mathcal{L}) :*

1. A est \mathcal{T} -pardonnante sur les solutions $\lambda\Psi$, \mathcal{L} -parcimonieuses.
2. A_r^λ est \mathcal{T}' -pardonnante.
3. $\text{Im}A_r^\lambda$ a la \mathcal{T}' -PNC (ou \mathcal{T} , \mathcal{L} -PNC).

Démonstration. $2 \Leftrightarrow 3$: On utilise directement le théorème 6.1.2 (chapitre 6).

$3 \Rightarrow 1$: En utilisant le théorème 6.1.2, $\text{Im}A_r^\lambda$ a la \mathcal{T}' -PNC implique que A_r^λ est \mathcal{T}' -pardonnante, en particulier, pour tout $w_r = \begin{pmatrix} Au_0 + n.T \\ 0 \end{pmatrix}$ défini comme dans la remarque 7.1.2. Alors, le problème (7.3) a u_0 comme unique solution pour tout u_0, \mathcal{T}' ce qui est équivalent au pardon sur les solutions $\lambda\Psi$, \mathcal{L} -parcimonieuses.

$1 \Rightarrow 2$: Soit A une matrice \mathcal{T} -pardonnante sur les solutions $\lambda\Psi$, \mathcal{L} -parcimonieuses, $L \in \mathcal{L}$ et $T \in \mathcal{T}$. Soit $u_0 \in \mathbb{R}^N$, $n_1 \in \mathbb{R}^N$, $n_2 \in \mathbb{R}^{N'}$. Soit $u_1 = \Psi^{-1}(n_2.L)$. En utilisant le pardon de A sur les solutions $\lambda\Psi$, \mathcal{L} -parcimonieuses, u_1 est l'unique minimiseur de

$$u_1 = \underset{u}{\text{argmin}} \|Au - (Au_1 + n_1.T)\|_1 + \lambda\|\Psi u\|_1. \quad (7.9)$$

On effectue le changement de variable $u = v + u_1 - u_0$. Alors u_0 est l'unique minimiseur :

$$\begin{aligned} u_0 &= \underset{v}{\text{argmin}} \|Av - (Au_0 + n_1.T)\|_1 + \lambda\|\Psi(v + u_1 - u_0)\|_1 \\ u_0 &= \underset{v}{\text{argmin}} \|Av - (Au_0 + n_1.T)\|_1 + \lambda\|\Psi v - \Psi u_0 - n_2.L\|_1 \end{aligned} \quad (7.10)$$

On conclut que A_r^λ est \mathcal{T}' -pardonnante. \square

Remarque 7.1.5. *Si Ψ est une décomposition générique dans un dictionnaire redondant, seulement la condition suffisante $3 \Leftrightarrow 2 \Rightarrow 1$ reste valide.*

Remarque 7.1.6. *D'après ce théorème, le pardon de A_r^λ est garanti si A est \mathcal{T} -pardonnante et Ψ est \mathcal{L} -pardonnante.*

Une application de ce théorème est le comportement de la régularisation lorsque $\lambda \rightarrow 0$.

Proposition 7.1.1. *Soit A une matrice \mathcal{T} -pardonnante alors, pour λ suffisamment petit, la matrice A_r^λ est \mathcal{T} -pardonnante sur les solutions parcimonieuses pour n'importe quelle parcimonie.*

Démonstration. Soit $T' = \begin{pmatrix} T \\ L \end{pmatrix} \in \mathcal{T}'$. Soit $\lambda > 0$ et u avec $\|u\|_1 = 1$. Parce que $\text{Im}A$ a la K -PNC avec constante $\gamma < 1$, on a :

$$\begin{aligned} \|(Au).T\|_1 + \lambda\|(\Psi u).L\|_1 &< \gamma\|(Au).T^c\|_1 + \lambda\|(\Psi u).L\|_1 \\ \|(A_r^\lambda u).T'\|_1 &< \gamma\|(Au).T^c\|_1 + \lambda(\|(\Psi u).L\|_1 - \|(\Psi u).L^c\|_1) + \lambda\|(\Psi u).L^c\|_1 \end{aligned} \quad (7.11)$$

Si $(\|(\Psi u).L\|_1 - \|(\Psi u).L^c\|_1) < 0$, $\text{Im}A_r^\lambda$ a la T, L -PNC. Sinon, pour λ suffisamment petit et indépendant de u (Ψ une application linéaire en dimension finie), on a

$$\gamma\|(Au).T^c\|_1 + \lambda(\|(\Psi u).L\|_1 - \|(\Psi u).L^c\|_1) < \|(Au).T^c\|_1 \quad (7.12)$$

et $\text{Im}A_r^\lambda$ a la T, L -PNC. La conclusion découle du Théorème 7.1.2. \square

Avec la propriété de non concentration, on peut aussi rejeter des cas où la régularisation parcimonieuse n'améliorera pas le pardon.

Proposition 7.1.2. *Si Ψ n'a pas la \mathcal{L} -PNC. Alors pour λ assez grand, A_r^λ n'a pas la \mathcal{T}, \mathcal{L} -PNC.*

Démonstration. Soit T, L des supports dans \mathcal{T}, \mathcal{L} respectivement. Il y a un u tel que $\|(\Psi u).L\|_1 > \|\Psi u.L^c\|_1$. Soit $0 < a < \|(\Psi u).L\|_1 - \|\Psi u.L^c\|_1$. Alors,

$$\|(Au).T\|_1 + \lambda\|(\Psi u).L\|_1 > \|(Au).T\|_1 + \lambda a + \lambda\|(\Psi u).L^c\|_1. \quad (7.13)$$

Pour λ assez grand, $\|(Au).T\|_1 + \lambda a > \|(Au).T^c\|_1$ et

$$\|(Au).T\|_1 + \lambda\|(\Psi u).L\|_1 > \|(Au).T^c\|_1 + \lambda\|\Psi u.L^c\|_1 \quad (7.14)$$

et A_r^λ n'a pas la \mathcal{T}, \mathcal{L} -PNC. \square

7.1.5 La Restricted Isometry Property pour le pardon régularisé

La RIP fait l'hypothèse de supports de cardinal particulier. Pour étudier le problème régularisé et sa robustesse aux outliers, on a besoin de la RIP pour des supports de la forme

$T = \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix}$ avec cardinalité $|\tau_1| = K$ sur les N premières coordonnées et $|\tau_2| = S$ sur

les dernières N' coordonnées. On dit qu'un tel T a cardinal K, S . D'après la définition 6.1.2 (chapitre 6), un annihilateur B de A_r est de taille $(N + N' - M) \times (N + N')$. On écrit B_T la restriction d'une matrice B aux colonnes qui correspondent au support T ($B_T x = B(x_T)$) où $x_T = x.T$.

Définition 7.1.3. *B a la Restricted Isometry Property d'ordre K, S et constante $\delta > 0$ si pour tout $x \in \mathbb{R}^{N+N'}$, tout support T de cardinal K, S*

$$(1 - \delta)\|x_T\|_2 \leq \|B(x_T)\|_2 \leq (1 + \delta)\|x_T\|_2. \quad (7.15)$$

La démonstration usuelle [24, 29] montre que la RIP implique la propriété de non concentration, ce qui implique le pardon. Ici, on montre qu'il est important de distinguer la parcimonie du signal et celle du bruit. On utilise la démonstration la plus simple qui est la plus restrictive sur B afin d'obtenir le résultat qualitatif. De meilleures constantes sont atteignables avec des arguments plus compliqués.

Théorème 7.1.3. *Soit B un annihilateur de A_r . Si B a la RIP d'ordre $3K, 3S$ avec constante $\delta \leq \frac{\sqrt{2}-1}{\sqrt{2}+1}$, alors $\ker B$ a la PNC d'ordre K, S et A_r est K, S pardonnante.*

Démonstration. Ce qui suit est une adaptation de la preuve de [29]. On prouve la PNC de $\ker B$. Soit $x \in \ker B$. Soit T_0 le support de cardinal K , S qui sélectionne les K plus grandes valeurs absolues de x sur les N premières coordonnées et S plus grandes valeurs absolues de x sur les N' dernières coordonnées. Soit $T_1 \dots T_r$ la suite des supports de cardinal $2K$, $2S$ ayant les plus grands éléments suivants. Soit $x_0 = x_{T_0} + x_{T_1}$, alors $Bx_0 = -B(\sum_{i=2}^r x_{T_i})$ car $x \in \ker B$. Il est suffisant de montrer la PNC pour $T = T_0$ car il concentre le plus d'énergie de x . En utilisant l'inégalité de Cauchy-Schwartz (où l'équivalence des normes) $\|x_T\|_1 \leq K^{1/2}\|x_{T,1}\|_2 + S^{1/2}\|x_{T,2}\|_2$ où $x_{T,1}$ est x restreint aux N premières coordonnées, et $x_{T,2}$ est x restreint aux N' dernières coordonnées.

Parce que $x_{0,1}$ et $x_{0,2}$ sont $3K$, $3S$ -parcimonieux, on utilise la définition de x_0 suivit par l'hypothèse RIP et l'inégalité triangulaire :

$$\begin{aligned}
 \|x_T\|_2 &\leq K^{1/2}\|x_{T,1}\|_2 + S^{1/2}\|x_{T,2}\|_2 \\
 &\leq K^{1/2}\|x_{0,1}\|_2 + S^{1/2}\|x_{0,2}\|_2 \\
 &\leq (1 - \delta)^{-1}(K^{1/2}\|Bx_{0,1}\|_2 + S^{1/2}\|Bx_{0,2}\|_2) \\
 &\leq (1 - \delta)^{-1} \left(K^{1/2}\|B \left(\sum_{i=2}^r x_{T_i,1} \right)\|_2 \right. \\
 &\quad \left. + S^{1/2}\|B \left(\sum_{i=2}^r x_{T_i,2} \right)\|_2 \right) \\
 &\leq (1 - \delta)^{-1} \left(\sum_{i=2}^r (K^{1/2}\|Bx_{T_i,1}\|_2 + S^{1/2}\|Bx_{T_i,2}\|_2) \right)
 \end{aligned} \tag{7.16}$$

On utilise la RIP de nouveau :

$$\|x_T\|_2 \leq \frac{1 + \delta}{1 - \delta} \left(\sum_{i=2}^r (K^{1/2}\|x_{T_i,1}\|_2 + S^{1/2}\|x_{T_i,2}\|_2) \right) \tag{7.17}$$

Maintenant, on borne le terme de droite avec $\|x\|_1$. Soit $j \geq 1$ et y (respectivement y') une des premières N coordonnées de $x_{T_{j+1}}$ (respectivement x_{T_j}). Alors $|y| \leq |y'|$. Soit z (respectivement z') une des dernières N' coordonnées de $x_{T_{j+1}}$ (respectivement x_{T_j}). Alors $|z| \leq |z'|$. D'après cette observation,

$$\begin{aligned}
 2Ky &\leq \|x_{T_j,1}\|_1 \\
 2Sz &\leq \|x_{T_j,2}\|_1.
 \end{aligned} \tag{7.18}$$

On élève au carré et on somme sur les y et z :

$$\begin{aligned}
 2K\|x_{T_{j+1},1}\|_2^2 &\leq \|x_{T_j,1}\|_1^2 \\
 2S\|x_{T_{j+1},2}\|_2^2 &\leq \|x_{T_j,2}\|_1^2
 \end{aligned} \tag{7.19}$$

On prend la racine et on somme les inégalités :

$$\sqrt{2}(K^{1/2}\|x_{T_{j+1},1}\|_2 + S^{1/2}\|x_{T_{j+1},2}\|_2) \leq \|x_{T_j}\|_1. \tag{7.20}$$

On combine avec l'équation (7.17) :

$$\|x_T\|_1 \leq \frac{1 + \delta}{1 - \delta} \frac{1}{\sqrt{2}} \|x_{T^c}\|_1 \tag{7.21}$$

Ce qui implique la PNC si $\frac{1}{\sqrt{2}} \frac{1+\delta}{1-\delta} \leq 1$. Ce qui est équivalent à $\delta \leq \frac{\sqrt{2}-1}{\sqrt{2}+1}$. \square

Remarque 7.1.7. *Ce théorème pourrait être étendu directement aux matrices B ayant plus de 2 blocs, par exemple pour un signal corrompu par des outliers et parcimonieux dans une union de dictionnaires (ou union de bases).*

Condition suffisante pour la RIP des matrices annihilatrices

Ici, on montre avec notre construction particulière de matrice annihilatrice, que l'on peut trouver une condition suffisante pour la K, S -RIP de A_r^λ . Soit $B_r = \begin{pmatrix} B_1 & B_2 \end{pmatrix}$ une annihilatrice de A_r^λ , avec B_1 une matrice $(N + N' - M) \times N$ et B_2 une matrice $(N + N' - M) \times N'$. On se demande à quelle condition B_r a la RIP.

Définition 7.1.4 (Bornes d'opérateur sur les vecteurs parcimonieux). *B est borné par σ_m, σ_M sur les vecteurs K -parcimonieux si :*

$$\sigma_M = \sup_{x, |T| \leq K} \frac{\|Bx_T\|_2}{\|x_T\|_2} \quad (7.22)$$

$$\sigma_m = \inf_{x, |T| \leq K} \frac{\|Bx_T\|_2}{\|x_T\|_2} \quad (7.23)$$

Proposition 7.1.3 (Normalisation). *Si B est borné par σ_m, σ_M sur les vecteurs K, S -parcimonieux, on peut trouver η tel que ηB a la RIP de constante $\delta = \frac{\sigma_M - \sigma_m}{\sigma_M + \sigma_m}$. On appelle ηB , B normalisé.*

Démonstration. Soit $\eta = 2/(\sigma_M + \sigma_m)$ et $\delta = \frac{\sigma_M - \sigma_m}{\sigma_M + \sigma_m}$, alors $\eta\sigma_m = 1 - \delta$ et $\eta\sigma_M = 1 + \delta$. \square

La cohérence mutuelle entre B_1 et B_2 (comme dans [19]) est utile pour produire des conditions suffisantes de RIP déterministe et probabiliste [19]. Cette mesure de la compatibilité de deux dictionnaires peut prendre plusieurs formes. Tout d'abord, considérons deux familles de vecteurs orthonormaux. Ces deux familles concaténées forment une famille orthonormale si les vecteurs sont orthogonaux deux à deux. C'est-à-dire si la valeur absolue maximale du produit scalaire entre deux vecteurs de chaque famille est nulle. Il est donc naturel de quantifier la compatibilité de deux familles de vecteurs (dictionnaire) par cette valeur. Dans la littérature, on se rend compte que plusieurs mesures de compatibilité sont utilisées. Il se trouve qu'elles ont toutes un sens très proche.

- Cohérence mutuelle : pour une concaténation de dictionnaires D , on la définit par $\mu_c = \max_{i,j} | \langle d_i, d_j \rangle |$ où les d_i sont les lignes de D .
- Propriété d'orthogonalité restreinte : dans le cadre de la reconstruction parcimonieuse, il s'agit du maximum du produit scalaire de deux observations de deux vecteurs parcimonieux normalisés. C'est cette version que l'on utilise. Cette quantité est bornée par la définition précédente, qui englobe aussi la RIP. On a donc

$$\mu = \sup_{|T| \leq K, |L| \leq S} \|B_{1,T}^H B_{2,L}\|_{op}. \quad (7.24)$$

On a $\mu \leq \max_{i,j} | \langle b_{1,i}, b_{2,j} \rangle |$ et finalement $\mu \leq \max_{i,j} | \langle b_i, b_j \rangle |$

- Angle entre sous-espaces : dans [71] la propriété d'orthogonalité restreinte est interprétée en tant qu'angle minimal entre deux sous-espaces.

Proposition 7.1.4 (Condition suffisante pour la RIP des annihilateurs de A_r^λ). *Soit B_1 borné par $\sigma_{m_1}, \sigma_{M_1}$ sur les vecteurs K -parcimonieux et B_2 borné par $\sigma_{m_2}, \sigma_{M_2}$ sur les vecteurs S -parcimonieux. Soit $\sigma_m = \min(\sigma_{m_1}, \sigma_{m_2})$ et $\sigma_M = \max(\sigma_{M_1}, \sigma_{M_2})$. Soit $\mu = \sup_{|T| \leq K, |L| \leq S} \|B_{1,T}^H B_{2,L}\|_{op}$. Si $\mu < \sigma_m^2$, alors B_r normalisé a la RIP d'ordre K, S et constante $\delta' = \frac{\sqrt{\sigma_M^2 + \mu} - \sqrt{\sigma_m^2 - \mu}}{\sqrt{\sigma_M^2 + \mu} + \sqrt{\sigma_m^2 - \mu}} < 1$.*

Démonstration. Soit $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ (x_1 a taille N) un vecteur de support T de cardinal K, S .

$$\begin{aligned} \|Bx\|_2^2 &= \|B_1 x_1 + B_2 x_2\|_2^2 \\ &= \|B_1 x_1\|_2^2 + \|B_2 x_2\|_2^2 + 2 \langle B_1 x_1, B_2 x_2 \rangle \end{aligned} \quad (7.25)$$

La cohérence mutuelle μ entre B_1 et B_2 joue un rôle important dans cette équation :

$$|\langle B_1 x_1, B_2 x_2 \rangle| \leq \mu \|x_1\|_2 \|x_2\|_2 \leq \frac{\mu}{2} \|x\|_2^2 \quad (7.26)$$

car $ab \leq (a^2 + b^2)/2$. Les 2 premiers termes sont contrôlés par les bornes de B_1 et B_2 . On a :

$$\begin{aligned} \min(\sigma_{m_1}, \sigma_{m_2})^2 - \mu &\leq \frac{\|Bx\|_2^2}{\|x\|_2^2} \leq \max(\sigma_{M_1}, \sigma_{M_2})^2 + \mu \\ \sqrt{\sigma_m^2 - \mu} &\leq \frac{\|Bx\|_2}{\|x\|_2} \leq \sqrt{\sigma_M^2 + \mu} \end{aligned} \quad (7.27)$$

La conclusion est une conséquence de la proposition 7.1.3. \square

Le choix de l'annihilateur

La qualité de la RIP dépend directement du conditionnement des sous-matrices de B . Le choix de B tel que $\ker B = \text{Im} A$ et tel que B a la RIP de constante optimale, est un problème ouvert. Une autre question peut même se poser. Étant donné une matrice d'observation sous-déterminée C , C peut très bien ne pas être la meilleure candidate pour vérifier la RIP. C'est-à-dire qu'il peut exister une matrice de même noyau ayant de meilleures constantes de RIP, ce qui garantit une meilleure capacité de reconstruction parcimonieuse. En revanche, comme on l'a vu au chapitre précédent, on peut construire une matrice B à partir de A en prenant la projection orthogonale sur le supplémentaire orthogonal de l'image de A ($P_{(\text{Im} A)^\perp}$), et restreindre celui-ci à son image. Alors, B est un annihilateur de A .

De plus, si on se place dans le cas où Ψ est inversible, on peut construire une famille d'annihilateurs de A_r en utilisant un annihilateur de A . Avec Ψ inversible, on a $N' = M$. Alors B est de taille $N \times N + M$, B_1 est de taille $N \times N$ et B_2 de taille $N \times M$.

Proposition 7.1.5. *Soit $\epsilon \geq 0$. Soit $\Phi = \Psi^{-1}$. Soit $B_r = \begin{pmatrix} B_1 & -\lambda^{-1} A \Phi \end{pmatrix}$, avec $B_1 = \epsilon P_{(\text{Im} A)^\perp} + I$. Alors B_r est un annihilateur de A_r^λ .*

Démonstration. On remarque d'abord que $B_2 = -\lambda^{-1}B_1A\Phi$ implique $B_rA_r^\lambda = 0$. Alors $B_r = \begin{pmatrix} B_1, & -\lambda^{-1}B_1A\Phi \end{pmatrix}$ est suffisant. De plus, B_1 est de rang plein et $B_1A = A$. \square

Remarque 7.1.8. On a construit B_r à partir de $P_{(\text{Im}A)^\perp}$ (annihilateur de A). On a ajouté I pour s'assurer que B_1 est de rang plein. Une telle construction permet de comparer l'amélioration des constantes de RIP causée par la régularisation.

Dans la Section 7.2.3, on montre sur un exemple que différentes matrices annihilatrices donnent des constantes de RIP différentes.

À propos du choix de λ

Le paramètre λ représente le niveau de connaissance a priori que l'on veut imposer à la solution. La question du choix de λ dans des contextes différents (ex. : minimisation L^2, L^2 ; minimisation L^2, L^1) est un sujet de recherche actif dans le domaine du traitement des images. Une manière de faire ce choix est de construire une fonction objectif, généralement une fonction de risque, et de trouver le λ qui minimise ce risque [108]. Ici, le cas de reconstruction parfaite est étudié. En conséquence, 2 questions viennent naturellement :

- Pour un λ , quels couples K, S mènent à une reconstruction parfaite.
- Étant donné K, S , peut-on trouver un λ garantissant une reconstruction parfaite (c.-à-d. le problème est K, S -pardonnant pour un λ).

Idéalement, on devrait choisir λ pour optimiser la constante de PNC ou garantir la PRE. Cela associerait chaque couple K, S à une constante de PNC optimale pour un λ optimal. Si cette constante était plus petite que 1, la reconstruction serait alors possible. Le problème est qu'elle est difficile en général à déterminer, car les calculs font intervenir le conditionnement L^1 des sous-matrices de A_r^λ . Pour certains problèmes de traitement des images, il peut être possible d'estimer cette constante directement. Dans des cas simples, on peut étudier directement la PNC et la PRE en fonction de λ (Section 7.3).

Une première étape est d'optimiser la RIP qui va donner un conditionnement suffisant pour le K, S -pardon. On montre que l'on peut garantir la RIP pour un λ avec notre construction de matrice annihilatrice pour une certaine classe de matrices A_r^λ . Comme on ne connaît pas d'annihilateur optimal (en terme de RIP), ce résultat à une utilité pratique limitée, mais illustre comment les blocs de B interagissent. Les expériences montreront que tester directement la RIP d'annihilateurs donne le comportement qualitatif du pardon de matrices aléatoires.

On regarde maintenant comment, à partir de A_r^λ , on peut calculer la constante de RIP δ' pour une matrice B_r particulière. Ce qui permet de trouver une condition suffisante sur λ qui garantit le $K/3, S/3$ -pardon de A_r^λ .

Théorème 7.1.4. On considère la matrice régularisée A_r^λ et la matrice B_r correspondante (construite comme dans la proposition 7.1.5) avec $\epsilon = 0$: $B_r = (I, -\lambda^{-1}A\Phi)$. Soit $A' = A\Phi$. On suppose que A' est bornée par $\sigma_m(A')$, $\sigma_M(A')$ sur les vecteurs S -parcimonieux. Soit $\mu_1 = \sup_{|T| < K, |L| < S} \|I_T^H A'_L\|_2$.

Si $\mu_1^2 \leq \sigma_m(A')^2$, alors on peut trouver λ tel que $\frac{\mu_1}{\sigma_m(A')^2} \leq \lambda^{-1} \leq \frac{1}{\mu_1}$. Avec un tel λ , la matrice B_r normalisée a la RIP d'ordre K, S et constante $\delta' \leq 1$.

En particulier pour $\lambda^{-1} = 1/\sigma_m(A')$, on a $\delta' \leq \frac{\sqrt{\kappa^2 + \mu_1} - \sqrt{1 - \mu_1}}{\sqrt{\kappa^2 + \mu_1} + \sqrt{1 - \mu_1}}$ où $\kappa = \sigma_M(A')/\sigma_m(A')$.

Démonstration. Soit $\sigma_m = \min(\sigma_{m_1}, \sigma_{m_2})$, $\sigma_M = \max(\sigma_{M_1}, \sigma_{M_2})$ comme définit proposition 7.1.4. On a $B_1 = I$, $\sigma_{m_1} = \sigma_{M_1} = 1$, alors :

$$\sigma_m = \min(1, \lambda^{-1} \sigma_m(A')) \quad (7.28)$$

et

$$\sigma_M = \max(1, \lambda^{-1} \sigma_M(A')) \quad (7.29)$$

Pour remplir l'hypothèse de la proposition 7.1.4, on veut $\mu = \lambda^{-1} \mu_1 \leq \sigma_m^2$. C'est vrai si les inégalités suivantes sont vraies :

$$\begin{aligned} \lambda^{-1} \mu_1 &\leq 1 \\ \lambda^{-1} \mu_1 &\leq \lambda^{-2} \sigma_m(A')^2 \end{aligned} \quad (7.30)$$

ce qui est équivalent à :

$$\begin{aligned} \lambda^{-1} &\leq \frac{1}{\mu_1} \\ \frac{\mu_1}{\sigma_m(A')^2} &\leq \lambda^{-1} \end{aligned} \quad (7.31)$$

Un tel λ^{-1} existe si :

$$\mu_1^2 \leq \sigma_m(A')^2 \quad (7.32)$$

Alors, avec la proposition 7.1.4, B_r a la RIP d'ordre K, S et constante $\delta' \leq 1$.

Si $\lambda^{-1} = 1/\sigma_m(A')$, on a $\sigma_m = 1$, $\sigma_M = \kappa$ et $\mu = \mu_1$. En utilisant ces valeurs :

$$\delta' \leq \frac{\sqrt{\kappa^2 + \mu_1} - \sqrt{1 - \mu_1}}{\sqrt{\kappa^2 + \mu_1} + \sqrt{1 - \mu_1}} \quad (7.33)$$

□

Remarque 7.1.9. On a montré en utilisant une famille de B_r particulière que l'on peut des fois déterminer laquelle donne la RIP.

Remarque 7.1.10. La condition $\mu_1^2 < \sigma_m(A')^2$ est forte. Des matrices le vérifient : prendre

par exemple la matrice $B_2 = \begin{pmatrix} I/\sqrt{P} \\ \vdots \\ I/\sqrt{P} \end{pmatrix}$, formée de P blocs. On a alors $\kappa = 1$ et

$$\mu_1 = 1/\sqrt{P} \text{ et } \delta' \leq \frac{\sqrt{\sqrt{P}+1} - \sqrt{\sqrt{P}-1}}{\sqrt{\sqrt{P}+1} + \sqrt{\sqrt{P}-1}}. \text{ Pour } P = 10, \text{ on a } \delta' \leq \frac{\sqrt{2}-1}{\sqrt{2}+1}$$

En utilisant la valeur explicite de λ du théorème, les hypothèses demandent que A' ait des sous matrices bien conditionnées, et une petite norme (faible cohérence mutuelle entre A et Φ) si $\kappa = 1.25$ et $\mu_1 = 0.1$, $\delta' \leq \frac{\sqrt{2}-1}{\sqrt{2}+1}$.

7.2 Le cas des matrices aléatoires

On étudie le cas de la régularisation convexe parcimonieuse et robuste pour les matrices aléatoires, ce qui permet de faire le lien avec la littérature. On suppose d'abord que $A_r^\lambda = \begin{pmatrix} A \\ \lambda \Psi \end{pmatrix}$ où A et Ψ sont des matrices $N \times M$ et $N' \times M$ gaussiennes aléatoires, \mathcal{T} est l'ensemble des supports $T \subset \mathbb{R}^N$ avec $|T| = K$ et $\mathcal{L} \subset \mathbb{R}^{N'}$ l'ensemble des supports L avec $|L| = S$. Dans la Section 7.2.3, la partie régularisation Ψ sera l'identité.

7.2.1 Revues des bornes existantes pour $\lambda = 1$

Soit B_r un annihilateur de A_r^λ . Dans [24], Candès et Tao argumentent que l'on peut identifier B à une matrice aléatoire gaussienne. La matrice B est de taille $N - M \times N$ ($\dim \text{Im} A = \dim \ker B = M$). Alors B a la K -propriété de reconstruction parcimonieuse et A est K -pardonnante avec probabilité dominante (« overwhelming ») si :

$$K < C \frac{N - M}{\log \frac{N}{N - M}} \quad (7.34)$$

avec C une constante. On veut voir si cette condition évolue lorsque l'on ajoute de la régularisation. On utilise la proposition suivante :

Proposition 7.2.1. *Si A_r^λ est $K + S$ -pardonnante, alors A est \mathcal{T} -pardonnante sur les vecteurs $\lambda \Psi$, \mathcal{L} -parcimonieux.*

Démonstration. $\text{Im} A_r^\lambda$ a la $K + S$ -PNC. De plus, l'ensemble des supports de taille $K + S$ contient $(\mathcal{T}, \mathcal{L})$. Alors, $\text{Im} A_r^\lambda$ a la $(\mathcal{T}, \mathcal{L})$ -PNC et A est \mathcal{T} -pardonnante sur les vecteurs $\lambda \Psi$, \mathcal{L} -parcimonieux. \square

A_r^λ est de taille $N + N' \times M$. On regarde l'évolution de la borne.

Proposition 7.2.2. *A est \mathcal{T} -pardonnante sur les vecteurs $\lambda \Psi$, \mathcal{L} -parcimonieux (ou A_r^λ est K, S -pardonnante) avec probabilité dominante (overwhelming) si :*

$$K + S < C \frac{N + N' - M}{\log \frac{N + N'}{N + N' - M}} \quad (7.35)$$

Démonstration. On utilise l'équation (7.34) avec les dimensions de A_r^λ et on applique la proposition 7.2.1. \square

Cette étude est théorique et prend le même type de matrice aléatoire pour l'observation et la régularisation, ce qui n'est pas très réaliste. Cependant, elle permet de montrer un exemple où ajouter de la régularisation augmente la capacité de pardon. L'observation principale est que c'est la quantité $K + S$ qui contrôle le pouvoir de pardon. Lorsqu'il y a plus d'outliers, le signal doit être plus parcimonieux d'après cette borne. Ceci est confirmé par l'expérience de la figure 7.2, où on a généré des expériences de pardon aléatoires pour différentes valeurs de K, S . On montre la pire de 100 expériences en terme de PSNR en décibels pour la minimisation du problème (7.4) (en utilisant $\ell_1 - \text{MAGIC}$, le logiciel utilisé par [25]).

Cette revue a montré que $K + S$ est le facteur pertinent pour $\lambda = 1$. Dans ce cas, chaque ligne de A_r^λ a la même énergie. Vérifier la $K + S$ -RIP de manière non uniforme donne le même résultat que pour la K, S -RIP, car c'est la même expérience.

Dans la preuve de la RIP des matrices aléatoires de [12] (où la condition sur $K + S$ est un peu différente que la borne de l'équation (7.34) mais n'est pas asymptotique), une borne d'union de probabilités est utilisée. La différence des probabilités pour la $K + S$ et K, S RIP viendra juste du nombre de sous matrices dont on regarde le conditionnement. Cette preuve repose sur l'inégalité

$$\binom{N + N'}{K + S} \leq \left(e \frac{N + N'}{K + S} \right)^{K + S}. \quad (7.36)$$

Elle permet de déduire que la RIP n'est pas valable avec probabilité au plus

$$\left(e \frac{N + N'}{K + S} \right)^{K + S} a^{K + S} e^{-b(N + N' - M)} \quad (7.37)$$

où a et b sont des constantes. Cette relation donne directement que si

$$K + S < C \frac{N + N' - M}{\log\left(\frac{N + N'}{N + N' - M}\right) + 1} \quad (7.38)$$

alors la RIP est satisfaite avec grande probabilité. Maintenant supposons que le nombre de sous-systèmes du problème de K, S -pardon donne la K, S -RIP (avec la même preuve que Baraniuk). Dans ce cas, on aurait besoin de

$$K \log\left(\frac{N}{K}\right) + S \log\left(\frac{N'}{S}\right) < C'(N + N' - M). \quad (7.39)$$

On voit que l'on pourrait s'attendre à une zone de validité de la RIP de forme différente, si on vérifiait effectivement tous les sous-systèmes (RIP uniforme).

Lorsque l'on vérifie la RIP non-uniforme, c'est à dire que l'on fait l'expérience pour r expériences indépendantes. on peut utiliser une borne d'union de la même manière. La probabilité d'échec devient inférieure à

$$\left(a^{K + S} e^{-b(N + N' - M)} \right)^r = e^{(K + S)r \log(a) - br(N + N' - M)}. \quad (7.40)$$

A paramètre fixé, on retrouve donc bien la condition $K < \text{constante}$ et a fortiori $K + S < \text{constante}$ pour la RIP. Lorsque l'on va effectuer les expériences non-uniformes, il faut donc garder à l'esprit que les zones de pardon décrites par la RIP non uniforme sont celles de la RIP uniforme à une constante près.

7.2.2 K, S -pardon en fonction de λ

On génère des expériences de K, S -pardon dans la figure 7.2. On produit 100 expériences aléatoires utilisant des matrices aléatoire gaussienne A de taille 100×50 et Ψ de taille 50×50 . Le PSNR de la pire expérience est représenté pour chaque K, S . On observe que pour chaque λ , on a des contrainte de type $K + c(\lambda)S < C_0$ où $c(\lambda)$ est une constante qui dépend de λ . Qualitativement, pour des λ grands, plus d'outliers sont enlevés mais cela

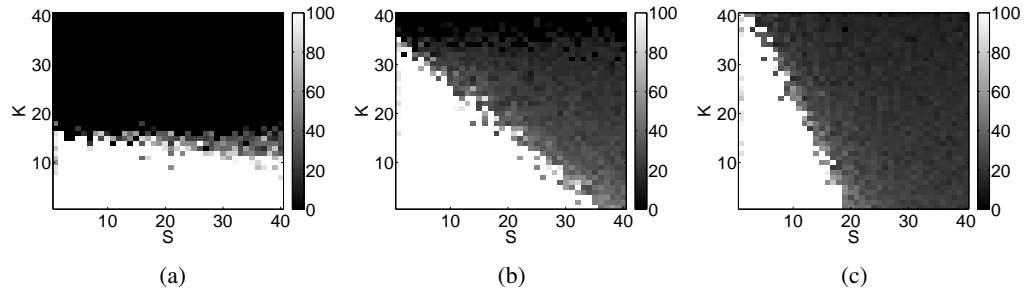


FIGURE 7.2. A aléatoire, Ψ aléatoire. Pire PSNR de reconstruction lorsque A_r^λ est de taille $N = 100$, $M = 50$ en fonction de la parcimonie du bruit et du signal K, S pour (a) $\lambda = 0.1$, (b) $\lambda = 1$ et (c) $\lambda = 2$.

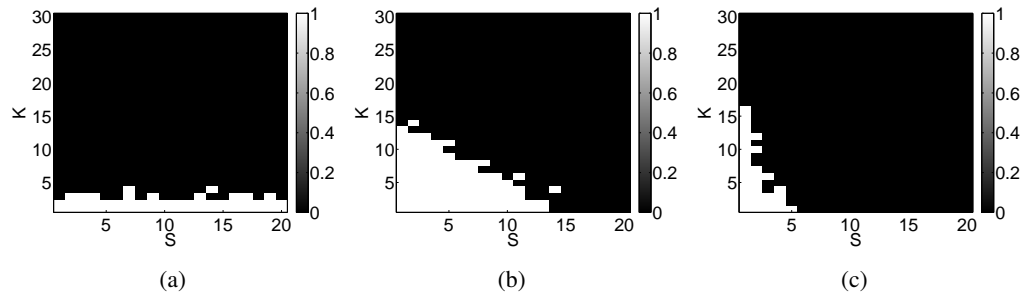


FIGURE 7.3. A aléatoire, Ψ aléatoire. Parcimonie du bruit et du signal K, S donnant la RIP avec A_r^λ est une observation aléatoire avec régularisation aléatoire de taille $N = 100$, $M = 50$ pour (a) $\lambda = 0.1$, (b) $\lambda = 1$ et (c) $\lambda = 2$.

n'est valable que pour les signaux très parcimonieux. Pour des λ plus petit, l'amélioration par rapport au cas non régularisé est moins forte mais demande moins de parcimonie.

Le résultat de ces expériences est en accord avec notre analyse théorique de la K, S -RIP. On montre expérimentalement que la K, S -RIP donne un moyen qualitatif de vérifier si la matrice est K, S -pardonnante. Soient A et Ψ des matrices aléatoires gaussiennes. On montre une estimation de la RIP de $P_{(\text{Im} A_r^\lambda)^\perp}$. Dans la figure 7.3, on génère le même nombre d'expérience, et on regarde si le pire conditionnement d'une sous-matrice par expérience a la constante de RIP requise. On montre les couples K, S où la RIP est vérifiée pour différentes valeurs de λ . L'union de tous les couples K, S montre où l'on peut avoir une reconstruction parfaite avec une régularisation parcimonieuse. La forme de la zone où les couples K, S donnent une reconstruction parfaite suit le même comportement que dans la figure 7.2. Comme attendu, un fossé quantitatif existe entre la caractérisation par la RIP et la capacité de pardon des expériences.

On a étudié le cas des matrices purement aléatoires pour faire le lien avec les résultats de la littérature. On étudie dans la section suivante un autre type de matrice où seulement l'observation est aléatoire.

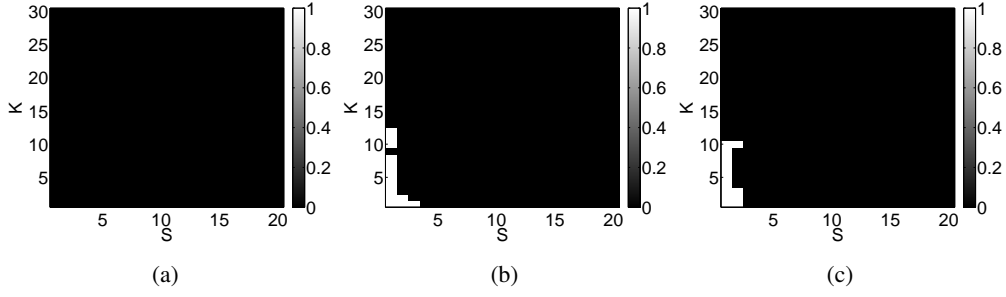


FIGURE 7.4. A aléatoire, $\Psi = I$. Parcimonie du bruit et du signal K, S donnant la K, S -RIP de $B_r = (I, -\lambda^{-1}A)$ avec $N = 100$, $M = 50$ pour (a) $\lambda = 0.1\sqrt{M}$, (b) $\lambda = \sqrt{M}$ et (c) $\lambda = 2\sqrt{M}$.

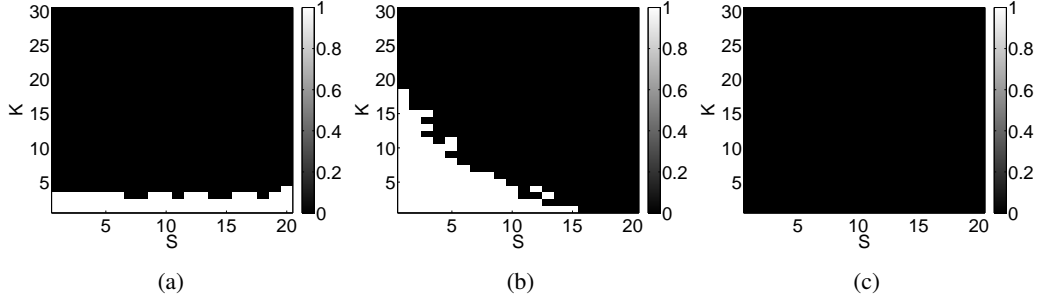


FIGURE 7.5. A aléatoire, $\Psi = I$. Même expérience que figure 7.4 avec un annihilateur différent. Parcimonie du bruit et du signal K, S donnant la K, S -RIP avec $N = 100$, $M = 50$ pour (a) $\lambda = 0.1\sqrt{M}$, (b) $\lambda = \sqrt{M}$ et (c) $\lambda = 2\sqrt{M}$.

7.2.3 Le cas d'observations aléatoires d'un signal parcimonieux

Ici on considère $A_r^\lambda = \begin{pmatrix} A \\ \lambda I \end{pmatrix}$. On effectue les mêmes expériences que dans la section

précédente. En utilisant la construction de B_r de la Section 7.1.5 pour $\epsilon = 0$, on a $B_1 = I$ et $B_2 = -\lambda^{-1}A$. Pour chaque λ , on regarde les couples K, S , pour lesquels, B_r a la RIP figure 7.4. Sur la figure 7.5, on fait la même expérience en utilisant $P_{(\text{Im} A_r^\lambda)^\perp}$. Cela illustre le fait que différentes matrices ayant le même noyau mènent à des RIP différentes. En conclusion, la RIP doit seulement être utilisée comme une mesure qualitative de la façon dont la reconstruction se comporte, et les annihilateurs doivent être choisis avec précaution.

Figure 7.6, l'expérience de reconstruction est faite en utilisant A_r^λ . Les variations de λ donnent le même comportement des zones donnant le K, S -pardon que dans le cas précédent.

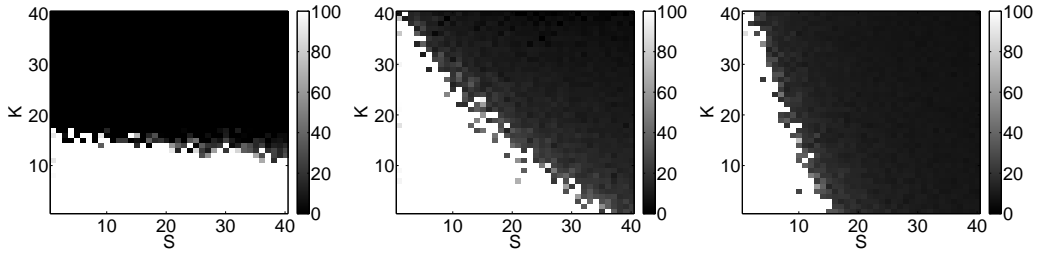


FIGURE 7.6. A aléatoire, $\Psi = I$. Pire PSNR de reconstruction en fonction de la parcimonie du bruit et du signal K, S , pour A_r^λ avec $N = 100$, $M = 50$ pour (a) $\lambda = 0.1\sqrt{M}$, (b) $\lambda = \sqrt{M}$ et (c) $\lambda = 2\sqrt{M}$

7.2.4 Conclusion sur les matrices aléatoires

La RIP donne un outils qualitatif pour déterminer la capacité de robustesse des problèmes décrits précédemment. Deux pistes à suivre sont apparues.

1. Déterminer la matrice annihilatrice optimale. Si on veut utiliser la RIP, trouver la matrice annihilatrice optimale peut contribuer à améliorer les constante théoriques.
2. Calcul analytique de la zone de pardon garantie (constante $c(\lambda)$). En étudiant plus profondément les preuves de la RIP des matrices aléatoires, il devrait être possible de déterminer le rôle de λ .

7.3 Robustesse de la super-resolution L^1 -TV

Dans cette section, on montre dans un cas particulier de traitement multi image (débruitage), que pour les supports ayant une valeur fixe de parcimonie K, S , la régularisation TV n'améliore pas le pardon. Cependant, si on fait l'hypothèse que les supports du gradient du signal et du bruit sont disjoints, on montre que le pardon est amélioré. On utilise ce résultat pour un problème de super-résolution particulier que l'on résout plus facilement en utilisant une minimisation $L^1 - TV$. Ces considérations mènent au calcul de la valeur optimale du paramètre de régularisation pour le débruitage multi-signal $L^1 - TV$ 1D et le problème de super-résolution. On finit par une expérimentation en 2D qui montre des résultats proche des conclusions théorique du cas 1D.

7.3.1 Débruitage multi-image

Le débruitage multi-image est le procédé utilisant différentes images d'une même scène pour produire une image avec un niveau de bruit réduit. On considère ici le problème simplifié de la robustesse aux outliers à partir d'une collection d'images recalées (pas de mouvement).

On utilise le modèle d'acquisition multi-image du chapitre 2 avec sous échantillonnage de facteur $M = 1$ sans mouvement ($Q_i = Id$). La minimisation L^1 de l'équation 6.1 mène à la sélection de la médiane des observations de chaque pixel. Dans ce cas, on peut directement étudier la propriété de non concentration de A et A_r^λ . Comme c'est un cas particulier de super-resolution, cela va nous permettre de trouver certaines limites de la SR $L^1 - TV$.

Soit N impair. En 1D, A est de taille $Nl \times l$ et est composée de blocs $l \times l$ diagonaux avec des coefficients égaux à 1. Soit $T \in \mathcal{T}$,

Proposition 7.3.1. *Dans le cas du débruitage 1D, A est \mathcal{T} -pardonnante si et seulement si chaque pixel est contaminé au plus $N/2$ -fois.*

Démonstration. On vérifie la \mathcal{T} -PNC pour A pour les deux implications.

D'abord, si un pixel i est contaminé $K = \lfloor N/2 \rfloor + 1$ fois, on considère le problème où seulement ce pixel n'est pas à 0. Alors

$$\|(Au).T\|_1 = K|u_i| > \lfloor N/2 \rfloor |u_i| = \|(Au).T^c\|_1 \quad (7.41)$$

et la PNC n'est pas vérifiée. Il est nécessaire que chaque pixel ne soit pas contaminé plus de $\lfloor N/2 \rfloor$ fois pour que A soit \mathcal{T} -pardonnante.

Réciproquement, si chaque pixel est contaminé plus de $\lfloor N/2 \rfloor$ fois, on regarde le pire des cas où chaque pixel est contaminé exactement $K = \lfloor N/2 \rfloor$ fois. On a

$$\|(Au).T\|_1 = K \sum_i |u_i| < (\lfloor N/2 \rfloor + 1) \sum_i |u_i| = \|(Au).T^c\|_1 \quad (7.42)$$

Alors $\text{Im}A$ a la \mathcal{T} -PNC. \square

On considère maintenant Ψ la fonction de gradient discret : $(\Psi u)_1 = u_1$, $(\Psi u)_i = u_i - u_{i-1}$. Comme Ψ est inversible, pour chaque support de parcimonie L , on peut trouver u tel que :

$$\|(\Psi u).L\|_1 > \|(\Psi u).L^c\|_1 \quad (7.43)$$

Si on considère le pire cas pour la PNC de A_r^λ , on trouve que Ψ ne peut pas l'améliorer. Soit un support de parcimonie L , on peut trouver un signal u tel que $\|(\Psi u).L\|_1 - \|(\Psi u).L^c\|_1 > 0$. Avec un tel u , on devrait s'assurer que :

$$\begin{aligned} \|(Au).T\|_1 + \lambda \|(\Psi u).L\|_1 &= K \sum_{i=1,S} |u_i| + \lambda \|(\Psi u).L\|_1 \\ &< (N - K) \sum_{i=1,S} |u_i| + \lambda \|(\Psi u).L^c\|_1. \end{aligned} \quad (7.44)$$

Ce qui est équivalent à

$$\begin{aligned} K \sum_{i=1,S} |u_i| + \lambda (\|(\Psi u).L\|_1 - \|(\Psi u).L^c\|_1) \\ < (N - K) \sum_{i=1,S} |u_i|. \end{aligned} \quad (7.45)$$

Le membre de gauche est minimisé pour $\lambda = 0$. C'est l'inégalité la plus faible que l'on peut obtenir. En conséquence, dans le pire des cas, il vaut mieux éviter la régularisation TV. En pratique, les outliers peuvent ne pas contaminer les sauts dans le signal. Par exemple, un reflet du soleil sur une surface réfléchissante uniforme remplira cette condition.

Supposons que les supports des outliers ne coïncident jamais avec les sauts du signal : c'est-à-dire que si le gradient à une position est différent de 0, les 2 échantillons utilisés pour son calcul ne sont pas contaminés par des outliers.

En utilisant la PRE, on peut montrer que le système devient très pardonnant, on peut même déterminer un λ optimal.

Proposition 7.3.2. *Soit \mathcal{T}' des supports de parcimonie de bruit et de signal, tel que le bruit et les sauts du signal sont disjoints. Si au plus K outliers contaminent chaque pixel et $\lambda > 0$ alors si $(2K - N)/2 < \lambda < N/2$, A_r^λ est \mathcal{T}' -pardonnante.*

Démonstration. On cherche un vecteur $v \in \mathbb{R}^{(N+1)l}$ qui vérifie les propriétés du théorème 7.1.1. On indexe les valeurs de v avec les indices $(v_{j,k})_{j=1..l, k=1..N+1}$. L'indice j représente le numéro de pixel. L'indice k représente le numéro de l'image acquise pour $k \leq N$ and le gradient de l'image pour $k = N + 1$. Soit $T' = (T, L)$ un support de \mathcal{T}' . Soit v tel que $v.T = \text{sign}(n).T$ et $v.L = \text{sign}(\Psi u_0).L$. On impose $(A_r^\lambda)^H v = 0$, ce qui se traduit pour chaque pixel (valeurs indexées par j de $(A_r^\lambda)^H v$) :

$$\sum_{k=0, N-1} v_{j,k} + \lambda(v_{j,N+1} - v_{j+1,N+1}) = 0 \quad (7.46)$$

On cherche maintenant à déterminer comment remplir la condition sur la norme infinie. Étant donné un pixel, on peut distinguer 2 cas :

- le pixel j n'est pas contaminé par des outliers. Alors, pour $k = 1..N$ chaque variable $v_{j,k}$ est libre et on peut choisir $v_{j,k} = -\lambda(v_{j,N+1} - v_{j+1,N+1})/N$. Comme le maximum de la norme L^1 du membre de droite est $2\lambda/N$, on garantit $\|v.T'^c\|_\infty < 1$ si $\lambda < N/2$.
- le pixel j est contaminé par des outliers. Alors on veut $\sum_{k \in T} v_{j,k} = -\sum_{k \in T^c} v_{i,k} - \lambda(v_{j,N+1} - v_{j+1,N+1})$, avec les variables du membre de droite qui sont libres. Comme la norme L^1 maximale du membre de gauche est K , on peut trouver un v correct avec $\|v.T'^c\|_\infty < 1$ si $K < N - K + 2\lambda$, ce que l'on réécrit $\lambda > (2K - N)/2$. □

La même proposition est vraie en 2D avec $(2K - N)/4 < \lambda < N/4$.

Remarque 7.3.1. Ici, \mathcal{T}' n'est pas la concaténation de 2 ensemble de supports de bruit et de signal.

Le résultat des expériences autorisant des superpositions entre les outliers et le gradient de signal est montré figure 7.7. Ils montrent que lorsque l'on permet ces superpositions, le débruitage $L^1 - TV$ n'est pas pardonnant.

Sur la figure 7.8, on montre que sans superposition entre bruit et saut de signal et avec $(N - 2)/2 < \lambda < N/2$, alors A_r^λ est $N - 1$ -pardonnante, pour $N = 7$. En d'autres mots, seulement une observation non contaminée par pixel est nécessaire pour avoir reconstruction parfaite.

7.3.2 Super-résolution

Les limites de la SR $L^1 - TV$ pour la robustesse aux outliers

Considérons que A est une matrice de SR translationnelle 1D et Ψ est le gradient discret (identique à la section précédente). L'étude du pire des cas dans le cas du débruitage montre que sans hypothèse précise sur le support du signal et des outliers, la régularisation TV n'améliore pas le pardon. Sur la figure 7.9, on montre 30 expériences de pardon de A et A_r^λ où A ne pardonne pas. On observe le compromis entre le nombre d'outliers et la parcimonie du signal pour une reconstruction parfaite.

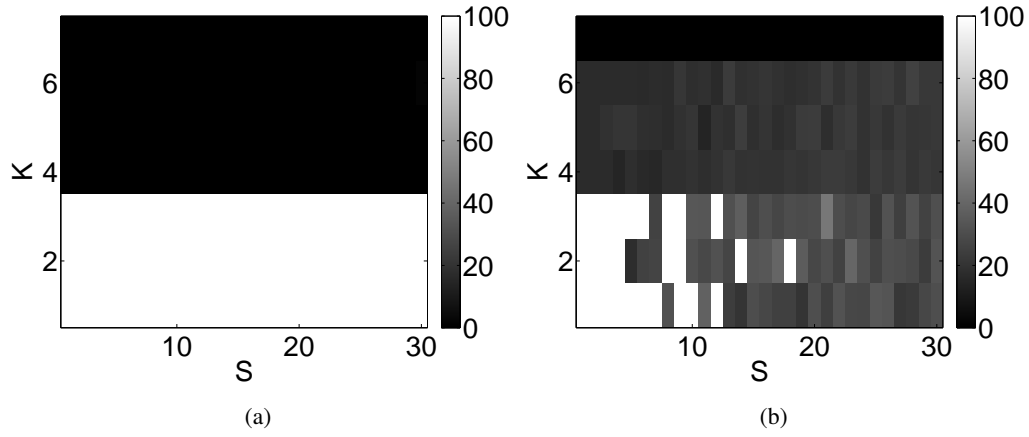


FIGURE 7.7. A observation multi-image, Ψ gradient. Avec recouvrement. Pire PSNR de la reconstruction en fonction de la parcimonie du signal et du bruit K, S lorsque les outliers peuvent recouvrir les sauts du signal avec $N = 7$ and $l = 50$ et (a) $\lambda = 0$, (b) $\lambda = 0.95N/2$

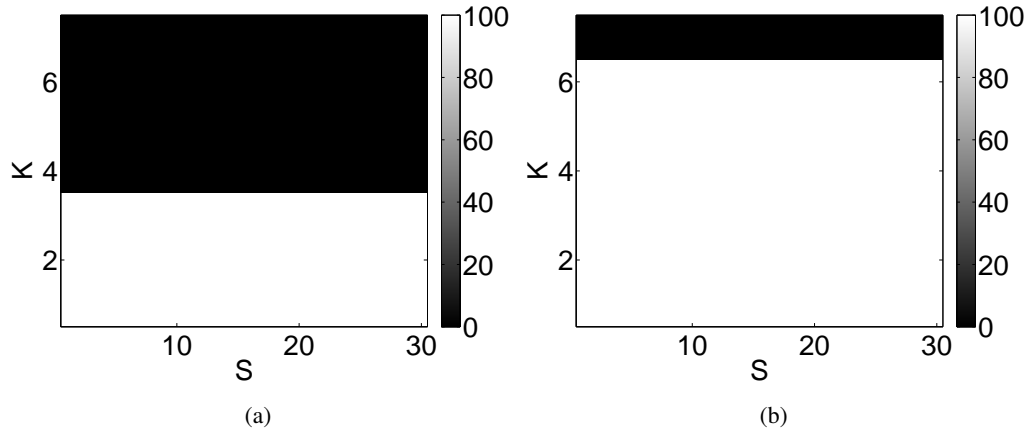


FIGURE 7.8. A = observation multi-image, Ψ = gradient. Pas de recouvrement Pire PSNR de la reconstruction en fonction de la parcimonie du signal et du bruit K, S lorsque les outliers ne peuvent pas recouvrir les sauts du signal avec $N = 7$ and $l = 50$ et (a) $\lambda = 0$, (b) $\lambda = 0.95N/2$

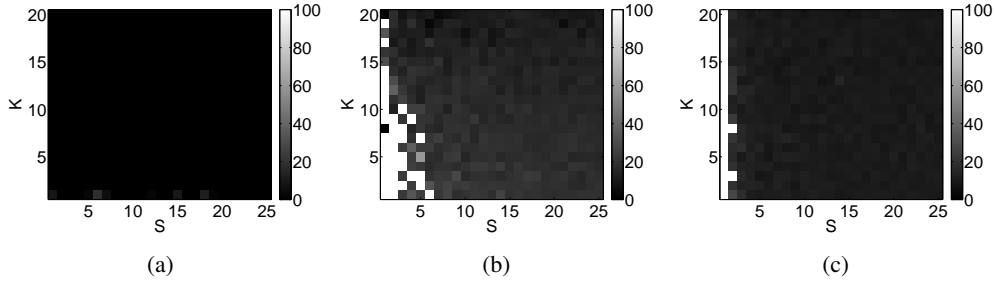


FIGURE 7.9. $A =$ **opérateur SR 1D**, $\Psi =$ **gradient**. Pire PSNR de la reconstruction avec régularisation TV en fonction de la parcimonie du signal et du bruit $N = 10$, $M = 2$ et $l = 25$ pour (a) $\lambda = 0$, (b) $\lambda = 1$ et (c) $\lambda = 10$.

La conclusion principale est que mettre une hypothèse de régularité permet d'être robuste à plus d'outliers, mais que l'amélioration est limitée comparée à la force de la supposition sur le signal. De plus ces expériences, ne donnent pas de garantie de pardon, car leur nombre est limité. Si on augmente le nombre d'expériences la zone de pardon diminue.

De la même manière que pour le débruitage multi-image, on veut savoir si l'hypothèse sur la superposition des supports apporte une amélioration dans le cas de la SR. Le principal problème que l'on rencontre pour étendre le résultat sur le débruitage est le suivant : les ligne de A sont des sinus cardinaux. Leur support n'est pas concentré. Il est donc difficile de définir une notion de recouvrement. En conséquence, la contamination par l'outlier va concerner tous les pixels de l'image HR.

La SR L^1 de réponse impulsionnelle finie bénéficie de la régularisation TV

Certaines méthodes de SR font l'hypothèse d'un opérateur de mouvement de longueur finie, un tel modèle peut être trouvé dans [26], et donne des résultats visuels probants. On propose de remplacer le mouvement par convolution avec un sinus cardinal, par une convolution avec une gaussienne tronquée. Un tel modèle peut être justifié par la connaissance de la vraie PSF (que l'on trouve dans [35]). Ici, pour la SR 1D, A est défini par

$$A : \mathbb{R}^{Ml} \rightarrow (\mathbb{R}^l)^N \quad (7.47)$$

$$u \rightarrow (A_i u)_{i=1,N} = (SG_i u)_{i=1,N}$$

où G_i est la convolution par une Gaussienne décalée tronquée $(g_q)_{q=-p,p}$ de taille $2p+1$. On choisit le paramètre Gaussien a pour que $|g_p|$ soit petit. On suppose donc que G_i est 0 hors des $p+3$ premières diagonales (l'amplitude des translations de la SR peut être considérée plus petite que $M=2$ sans perte de généralité). On garde la même définition de Ψ . Pour éviter des calculs compliqués, on suppose aussi que les sauts de signal sont séparés par la taille de G_i .

Proposition 7.3.3. *Soit T' des supports de parcimonie tels que les outliers et les sauts de signal sont séparés par $p+1$ échantillons et que les sauts sont séparés par $2p+1$ pixels. On suppose qu'au plus K outliers contaminent chaque pixel. Soit $G_+ = \max_{i,j} \|G_{i,j}\|_1$ et $G_- = \min_{i,j} \|G_{i,j}\|_1$ où $G_{i,j}$ est la j -ème colonne de SG_i . Si $(K(G_+ + G_-) - NG_-)/2 < \lambda < NG_-/2$, A_r^λ est T' -pardonnante.*

Démonstration. On cherche $v \in \mathbb{R}^{Nl+Ml}$ vérifiant les propriétés du théorème 7.1.1. Soit $T' = (T, L)$ un support de \mathcal{T}' . Soit v tel que $v.T = \text{sign}(n).T$ et $v.L = \text{sign}(\Psi u_0).L$. On écrit $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ où v_1 est de taille Nl et v_2 est de taille Ml . On veut $(A_r^\lambda)^H v = 0$. Ce qui s'écrit pour chaque pixel HR (chaque colonne j de A_r^λ) :

$$A_j^H v_1 + \lambda \Psi_j^H v_2 = 0, \quad (7.48)$$

avec A_j la j -ème colonne de A et Ψ_j la j -ème colonne de Ψ . Pour un pixel, on distingue 2 cas :

- le pixel j n'est pas contaminé par des outliers. Alors chaque variable de v_1 est libre. On veut $A_j^H v_1 > 2\lambda$ car la valeur maximale de $|\Psi_j^H v_2|$ est 2. La restriction $|v_1|_\infty < 1$ impose

$$\lambda < \|A_j^H\|_1/2. \quad (7.49)$$

- le pixel j est contaminé. On veut :

$$A_j^H(v_1.T) = -A_j^H(v_2.T^c) - \lambda \Psi_j^H v_2 \quad (7.50)$$

avec le membre de droite composé de variables libres. De la même façon que précédemment, il est nécessaire que $\|A_{T|j}^H\|_1 < \|A_{T^c|j}^H\|_1 + 2\lambda$, ce qui est équivalent à

$$\lambda > (\|A_{T|j}^H\|_1 - \|A_{T^c|j}^H\|_1)/2. \quad (7.51)$$

Les conditions (7.49) et (7.51) sont suffisantes. Avec la définition de G_+ et G_- , on a les inégalités :

$$\begin{aligned} NG_- &< \|A_j^H\|_1 \\ (K(G_+ + G_-) - NG_-)/2 &> \|A_{T|j}^H\|_1 - \|A_{T^c|j}^H\|_1 \end{aligned} \quad (7.52)$$

La condition $(K(G_+ + G_-) - NG_-)/2 < \lambda < NG_-/2$ est donc suffisante pour la PRE. \square

Remarque 7.3.2. On peut trouver un λ si $K < K_{max} = N/(1+G_+/G_-)/2$. Typiquement, $G_+/G_- \sim 2$. Si des sauts de signal consécutifs sont autorisés, les constantes changeraient et nécessiteraient la norme L^1 de matrices inverses de restrictions de A . Estimer de telles normes de manière analytique est difficile.

Expérience pour la super-résolution 1D : Dans la figure 7.10, on montre l'influence de la régularisation TV pour la SR 1D de réponse impulsionnelle finie avec un zoom $M = 2$. L'image est de taille $l = 210$. Pour chaque K, S , on génère 100 expériences avec les hypothèses de la proposition 7.3.3, en contaminant le même pixel dans les 10 images BR dans une zone sans saut. On résout le problème régularisé en utilisant notre calcul de λ . On voit qu'après régularisation, le nombre de pixels contaminés peut être plus grand que la moitié du nombre d'images.

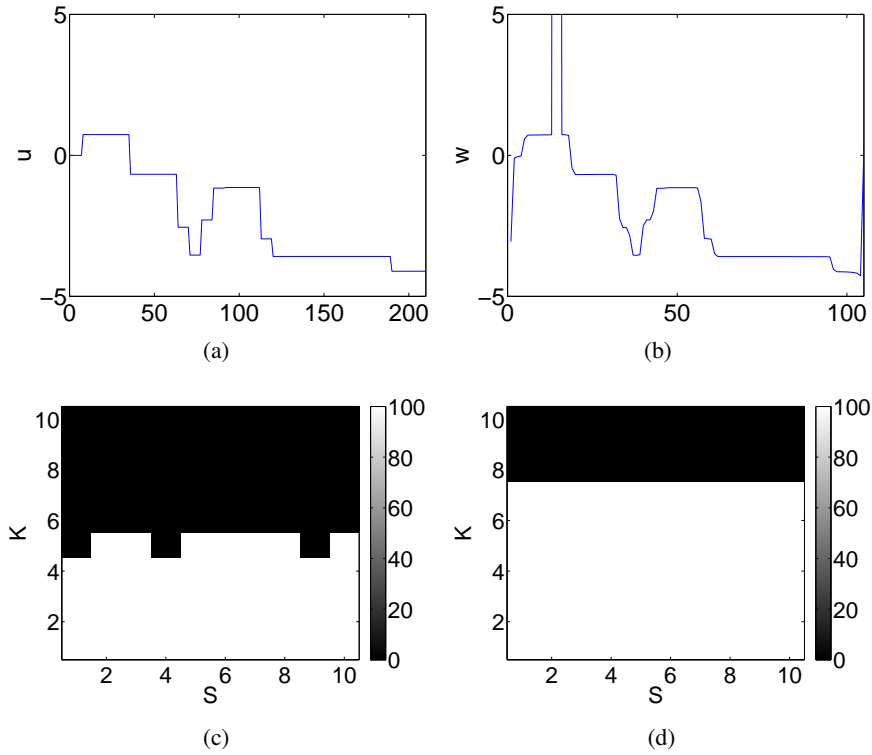


FIGURE 7.10. $A =$ opérateur SR 1D de réponse impulsionnelle finie, $\Psi =$ gradient. Pire PSNR de reconstruction pour différentes parcimonies de bruit et de signal K, S sans superposition ($N = 10, M = 2, l = 20$). (a) Exemple d’une image HR, (b) exemple d’une image BR contaminée, (c) SR sans régularisation (d) SR avec régularisation TV avec λ optimal.

Expériences pour la super-résolution 2D : On utilise le même modèle pour la super-résolution 2D (on utilise une gaussienne 2D). On utilise l’image “Shepp Logan phantom” dont le gradient est parcimonieux. Ψ est le gradient discret 2D. On effectue une super-résolution L^1 2D avec régularisation par variation totale avec $M = 2$. On utilise un algorithme de moindres carrés repondérés [31]. On contamine la même région de l’image dans les différentes images BR. Dans la figure 7.11, on montre l’image idéale, le niveau de parcimonie, une image BR sans bruit et une image BR contaminée.

D’après la proposition 7.3.3, on peut inférer qu’une plage de valeurs de λ rend le problème SR $L^1 - TV$ 2D pardonnant. On montre figure 7.12 le résultat de la SR L^1 2D sans régularisation et avec régularisation TV. Il y avait 6 images, $N_c = 2$ images contaminées (on rappelle que le nombre minimal d’images pour une reconstruction parfaite est de 4 pour $M = 2$). Alors que la super-résolution sans régularisation échoue, la régularisation TV donne une reconstruction parfaite. On montre aussi comment le PSNR se comporte en fonction de λ .

Une plage de valeurs de λ donne une reconstruction parfaite (de 0.1 à 0.8), en dehors on voit l’effet du sur-lissage (de 0.8 à 4) ou l’échec (petit λ).

Dans la figure 7.13, on utilise $N = 10$ et $N_c = 6$. le problème non régularisé échoue

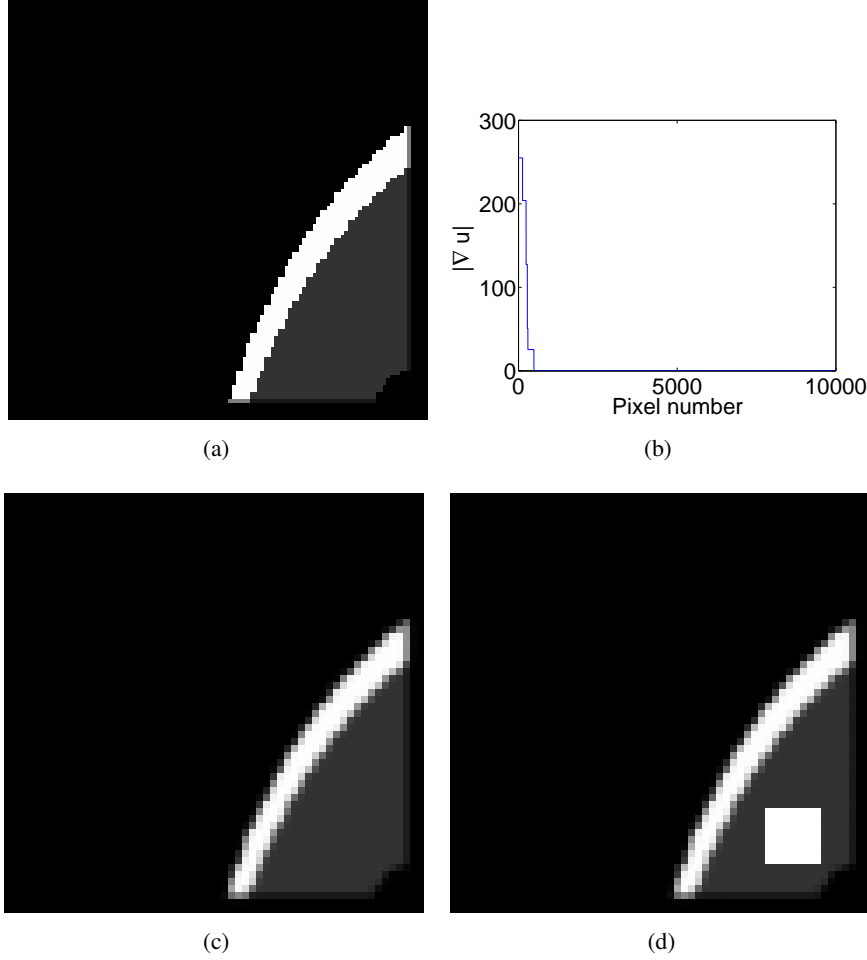


FIGURE 7.11. Le problème SR 2D. (a) Image HR, (b) Valeurs triées du gradient 2D discret (parcimonie du gradient), (c) une image BR propre (d) une image BR contaminée.

complètement. La régularisation TV mène à une reconstruction parfaite. Dans cette expérience, seulement le nombre critique d'images $M = 4$ est nécessaire pour retrouver l'image HR et il y a moins d'images propres que d'images contaminées. Le comportement de la reconstruction en fonction de λ donne un comportement similaire au précédent. Une plage de λ (de 0.9 à 1.05) donne une reconstruction parfaite.

Remarque 7.3.3. *On a supposé que seulement du bruit parcimonieux contaminait le signal. Cette hypothèse a été faite pour maintenir claire l'exposition des concepts principaux. Comme la théorie est une extension directe de la reconstruction parcimonieuse, la stabilité par rapport au bruit additif (ex. : bruit gaussien) devrait être obtenue en utilisant les résultats correspondants de la littérature [23, 29], que l'on a présentés dans la section 6.1.4.*

Autres schémas de régularisation, détection d'outliers et méthodes L^0

Comme noté précédemment, une méthode de SR par *sinc* ne bénéficie pas beaucoup de la régularisation TV. Deux voies possibles pourraient être explorées.

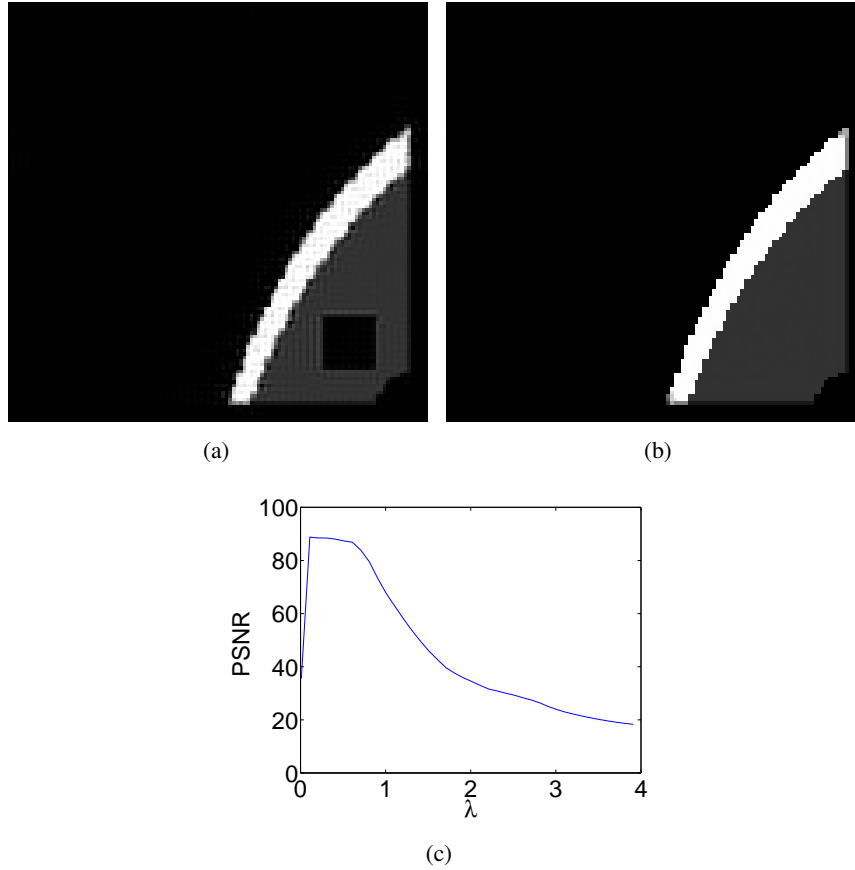


FIGURE 7.12. SR 2D résultats, expérience 1. (a) reconstruction sans régularisation, PSNR=25.49dB, (b) reconstruction avec régularisation PSNR=88.27dB, (c) PSNR en fonction de λ .

Premièrement, un moyen de s'assurer de la robustesse du problème est de s'assurer que la matrice surdéterminée Ψ est \mathcal{L} -pardonnante avec une bonne constante de PNC. Les approches non locales ou par dictionnaire pour la modélisation des images peuvent mener à un Ψ étant une projection dans un espace de grande dimension où le signal est parcimonieux. Si cette projection à la PNC relative aux supports de parcimonie du signal, alors le pardon sera amélioré pour tout type d'observation (encodage universel).

Deuxièmement, on peut faire les observations suivantes : les outliers très énergétiques et concentrés sont faciles à détecter. Alors que la minimisation $L^1 - TV$ sera incapable de les enlever dans certains cas critiques (*p. ex.* pour une super-résolution 2D avec $N = M^2 + 1$), si seulement un outlier contamine un saut, la détection par seuillage suivie de l'oubli de l'équation correspondante donnera un meilleur résultat que la minimisation par régularisation TV présentée dans ce chapitre.

Enfin, on peut montrer que la minimisation L^0 peut augmenter le pouvoir de pardon des outliers car elle cherche à maximiser le nombre d'équations satisfaites. L'apport en terme de pardon d'une telle méthode peut être important. Le cas précédent d'un outlier contaminant un saut dans un cas presque critique en est un exemple.

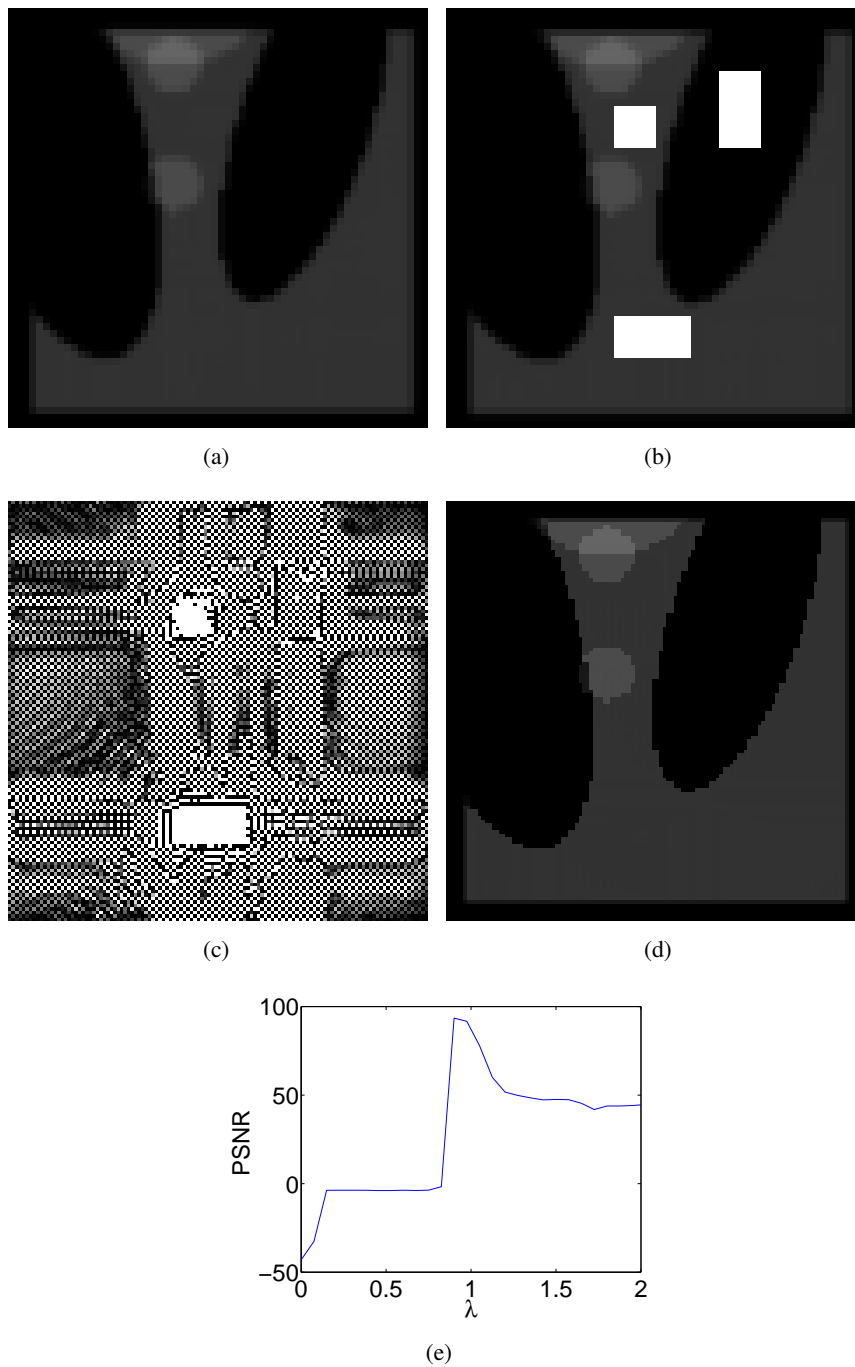


FIGURE 7.13. SR 2D résultats, expérience 2. (a) image HR, (b) une image BR contaminée, (c) reconstruction sans régularisation, PSNR=-46.84dB, (d) reconstruction avec régularisation PSNR=93.31dB, (e) PSNR en fonction de λ .

CHAPITRE 8

CONCLUSION

Le traitement multi-image est une manière de dépasser les limites physiques d'un appareil de photographie. L'objectif est de simuler un capteur plus performant en multipliant les acquisitions de photographies, et donc la quantité d'information capturée. On peut opposer à cette manière de faire, les techniques de régularisation, où c'est par une connaissance a priori de la scène à capturer, que l'on apporte l'information manquante. Notre but a été de déterminer quand la régularisation doit intervenir ou pas, c'est-à-dire les conditions pour lesquelles l'information manque ou pas. Dans le cas où l'information manque, on a essayé de déterminer quel était l'apport de la régularisation.

Ces travaux ont permis de déterminer des circonstances pour lesquelles aucune hypothèse de régularité n'est nécessaire pour retrouver la scène originale, lorsque celle-ci ne subit pas de modification pendant la prise des photographies. Des expériences ont montré que certaines de ces circonstances peuvent être réunies lors de situations réelles, c'est-à-dire que l'on peut effectivement prendre le nombre de photographies nécessaire à une bonne reconstruction de la scène. On a souvent pris des hypothèses pessimistes, il est donc possible que des paramètres d'acquisition moins contraignants donnent de bons résultats en pratique.

Dans le cas de défauts dans certaines images basse résolution (qui peuvent être des changements dans l'image), on a montré théoriquement que si la part d'images contaminées n'est pas trop grande, on peut retrouver la scène originale.

Il faut cependant garder à l'esprit qu'augmenter la quantité de données n'est pas toujours possible et qu'une régularisation est alors nécessaire. Les travaux que l'on a effectués sur le conditionnement local montre que la régularisation n'est pas forcément utile sur toute l'image dans le cas où le conditionnement global est mauvais. Si les données sont corrompues par des outliers qui ont une structure particulière et que l'on peut faire une hypothèse de type parcimonie, la robustesse est grandement améliorée.

D'un point de vue théorique, les travaux sur la robustesse ont permis de déterminer des outils utiles à l'étude du problème de super-résolution, et même plus généralement de traitement multi-image. Ces outils théoriques issus de la théorie de la reconstruction

	Régularisation peu/pas nécessaire
Pas de bruit	$N > M^2$
Bruit de puissance σ^2	N/σ^2 assez grand
N_c outliers	N/N_c assez grand

TABLE 8.1. Conditions sur le rapport signal / bruit pour une super-résolution bien informée

	Apport de la régularisation
Bruit d'énergie finie	Une régularisation TV locale peut préserver les zones bien conditionnées
N_c outliers	Avec des supports de parcimonie du signal et du bruit disjoints, N/N_c peut être petit, sinon apport nul.

TABLE 8.2. Apport de la régularisation

parcimonieuse et du *compressed sensing* ne sont pas forcément les plus utilisés : on a vu que les propriétés de non concentration et de reconstruction exacte sont plus pertinentes. Cette différence dans l'utilité de chaque outil s'explique par le fait que l'opérateur considéré a une structure particulière qui s'accorde mal des outils les plus populaires que sont la RIP et la cohérence mutuelle.

Perspectives

Dans ces travaux, de nombreuses perspectives ont été évoquées pour améliorer des points particuliers. On peut cependant dégager des directions un peu plus générales de recherche.

On a trouvé des conditions suffisantes pour que la super-résolution soit bien informée. Il serait intéressant de continuer à chercher des conditions plus larges qui garantissent une bonne reconstruction sans régularisation. Par exemple, il semble possible de considérer des mouvements plus généraux (homographies), d'introduire un flou inversible (estimé a priori) dans le modèle, et conserver les mêmes résultats.

Ce travail peut aussi servir de base pour la conception de systèmes de super-résolution. L'avantage d'avoir des algorithmes simples et une quantité de données suffisante, est que l'effort de conception d'un système de super-résolution repose sur l'amélioration de la puissance de calcul du matériel et l'amélioration des performances algorithmiques. Il faut aussi garder à l'esprit que les appareils photo possèdent un filtre anti-repliement physique. Il serait intéressant d'expérimenter avec des systèmes sans filtre, les possibilités en terme de gain de résolution pourraient être importantes.

On a seulement étudié en détail l'interaction de la régularisation TV avec le problème de super-résolution multi-image. Bien d'autres régularisations peuvent s'inscrire dans la théorie que l'on a décrite. On peut notamment citer les méthodes de régularisation à base de décomposition parcimonieuse dans des dictionnaires. Il serait intéressant d'étudier leur comportement en terme de robustesse pour la super-résolution, et plus particulièrement d'étudier si prendre en compte l'opérateur d'acquisition dans la construction du dictionnaire serait bénéfique à ces performances. Pour une telle étude, chercher à utiliser ou développer des bornes inférieures pour la norme L^1 de familles de vecteurs serait bénéfique.

Enfin, l'utilisation d'algorithmes de minimisation L^0 approchés semble être une tendance forte dans le domaine du traitement de l'image. Il est possible de montrer dans certains cas qu'ils améliorent les résultats de la minimisation L^1 . Étudier leurs performances

théoriques dans le cadre de la super-résolution et les comparer aux méthodes L^1 permettrait de valider leur utilité. Pour ce faire, on pourrait d'abord étudier le potentiel apport d'une minimisation L^0 exacte. Si le résultat est positif, il faudrait choisir le schéma de minimisation approché qui donne le meilleur compromis performance/temps d'exécution pour le traitement multi-image. Enfin, les remarques sur la structure de l'opérateur de super-résolution devraient jouer un rôle crucial pour la conception d'algorithmes approchant au mieux la minimisation L^0 .

BIBLIOGRAPHIE

- [1] C. AGUERREBERE, J. DELON, Y. GOUSSEAU et P. MUSÉ : Best algorithms for HDR image generation. a study of performance bounds. *SIAM Journal on Imaging Sciences*, 7(1):1–34, 2014.
- [2] C. AGUERREBERE, J. DELON, Y. GOUSSEAU et P. MUSÉ : Study of the digital camera acquisition process and statistical modeling of the sensor raw data. *Technical report* <http://hal.archives-ouvertes.fr/hal-00733538>, sept. 2012.
- [3] C. AGUERREBERE, J. DELON, Y. GOUSSEAU et P. MUSE : Simultaneous hdr image reconstruction and denoising for dynamic scenes. In *Computational Photography (ICCP), 2013 IEEE International Conference on*, p. 1–11, April 2013.
- [4] N. A. AHUJA et N. K. BOSE : Multidimensional Generalized Sampling Theorem for wavelet Based Image Superresolution. In *Image Processing, 2006 IEEE International Conference on*, p. 1589–1592. IEEE, oct. 2006.
- [5] A. AJDARI RAD, L. MEYLAN, P. VANDEWALLE et S. SÜSTRUNK : Multidimensional image enhancement from a set of unregistered differently exposed images. In *Proc. IS&T/SPIE Electronic Imaging : Computational Imaging V*, vol. 6498, 2007.
- [6] A. ALMANSA, J. CARON et S. DURAND : Deblurring of irregularly sampled images by TV regularization in a spline space. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, p. 1181–1184. IEEE, sept. 2010.
- [7] A. ALMANSA, S. DURAND et B. ROUGÉ : Measuring and Improving Image Resolution by Adaptation of the Reciprocal Cell. *Journal of Mathematical Imaging and Vision*, 21(3):235–279, nov. 2004.
- [8] ANTONI BUADES, BARTOMEU COLL, JEAN-MICHEL MOREL, CATALINA SBERT : Self-Similarity Driven Demosaicking. *Image Processing On Line*, 2011.
- [9] M. ARIGOVINDAN, M. SUHLING, P. HUNZIKER et M. UNSER : Variational image reconstruction from arbitrarily spaced samples : a fast multiresolution spline solution. *Image Processing, IEEE Transactions on*, 14(4):450–460, 2005.
- [10] S. D. BABACAN, R. MOLINA et A. K. KATSAGGELOS : Variational Bayesian Super Resolution. *Image Processing, IEEE Transactions on*, 20(4):984–999, avr. 2011.
- [11] S. BAKER et T. KANADE : Limits on super-resolution and how to break them. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(9):1167–1183, sept. 2002.
- [12] R. BARANIUK, M. DAVENPORT, R. DEVORE et M. WAKIN : A Simple Proof of the Restricted Isometry Property for Random Matrices. *Constructive Approximation*, 28(3):253–263, déc. 2008.
- [13] R. F. BASS et K. GROCHENIG : Random sampling of multivariate trigonometric polynomials. *SIAM journal on mathematical analysis*, 36(3):773–795, 2005.

- [14] T. BENGTTSSON, I. Y. GU, M. VIBERG et K. LINDSTRÖM : Regularized optimization for joint super-resolution and high dynamic range image reconstruction in a perceptually uniform domain. *In Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, p. 1097–1100. IEEE, mars 2012.
- [15] L. BENOIT, J. MAIRAL, F. BACH et J. PONCE : Sparse image representation with epitomes. *In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, p. 2913–2920. IEEE, juin 2011.
- [16] P. BHAT, K. C. ZHENG, N. SNAVELY, A. AGARWALA, M. AGRAWALA, M. F. COHEN et B. CURLESS : Piecewise Image Registration in the Presence of Multiple Large Motions. *In Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2, p. 2491–2497. IEEE, 2006.
- [17] S. BOYD et L. VANDENBERGHE : *Convex optimization* - Google Livres.
- [18] E. O. BRIGHAM : *The Fast Fourier Transform*. 1974.
- [19] E. CANDÈS et J. ROMBERG : Sparsity and incoherence in compressive sampling. *Inverse Problems*, 23(3):969–985, juin 2007.
- [20] E. J. CANDÈS et C. FERNANDEZ-GRANDA : Towards a Mathematical Theory of Super-resolution. *Comm. Pure Appl. Math.*, p. n/a, jan. 2013.
- [21] E. J. CANDES et Y. PLAN : Matrix Completion With Noise. *Proceedings of the IEEE*, 98(6):925–936, juin 2010.
- [22] E. J. CANDES, J. ROMBERG et T. TAO : Robust uncertainty principles : exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, fév. 2006.
- [23] E. J. CANDÈS, J. K. ROMBERG et T. TAO : Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59(8):1207–1223, août 2006.
- [24] E. J. CANDES et T. TAO : Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203–4215, déc. 2005.
- [25] E. J. CANDES et T. TAO : Near-Optimal Signal Recovery From Random Projections : Universal Encoding Strategies ? *Information Theory, IEEE Transactions on*, 52(12): 5406–5425, déc. 2006.
- [26] F. CHAMPAGNAT, G. LE BESNERAIS et C. KULCSÁR : Statistical performance modeling for superresolution : a discrete data-continuous reconstruction framework. *J. Opt. Soc. Am. A*, 26(7):1730–1746, juil. 2009.
- [27] T. F. CHAN et P. MULET : On the Convergence of the Lagged Diffusivity Fixed Point Method in Total Variation Image Restoration. *SIAM J. Numer. Anal.*, 36(2):354–367, 1999.
- [28] J. CHOI, M. K. PARK et M. G. KANG : High dynamic range image reconstruction with spatial resolution enhancement. *Comput. J.*, 52(1):114–125, jan. 2009.
- [29] A. COHEN, W. DAHMEN et R. DEVORE : Compressed sensing and best k-term approximation. *J. Amer. Math. Soc.*, 22(1), 2009.

-
- [30] B. COULANGE : *Détection et correction de l'aliasing par extension du signal analytique aux images numériques*. Thèse de doctorat, 2011.
 - [31] I. DAUBECHIES, R. DEVORE, M. FORNASIER et C. S. GÜNTÜRK : Iteratively re-weighted least squares minimization for sparse recovery. *Comm. Pure Appl. Math.*, 63(1):1–38, 2010.
 - [32] G. DAVIS, S. MALLAT et M. AVELLANEDA : Adaptive greedy approximations. *Constructive Approximation*, 13(1):57–98, mars 1997.
 - [33] Y. de CASTRO et F. GAMBOA : Exact reconstruction using Beurling minimal extrapolation. *Journal of Mathematical Analysis and Applications*, 395(1):336–354, nov. 2012.
 - [34] P. E. DEBEVEC et J. MALIK : Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, p. 369–378, 1997.
 - [35] M. DELBRACIO, A. ALMANSA, J.-M. MOREL et P. MUSÉ : Subpixel Point Spread Function Estimation from Two Photographs at Different Distances. *SIAM Journal on Imaging Sciences*, 5(4):1234–1260, nov. 2012.
 - [36] M. DELBRACIO, P. MUSÉ, A. ALMANSA et J.-M. MOREL : The Non-parametric Sub-pixel Local Point Spread Function Estimation Is a Well Posed Problem. *International Journal of Computer Vision*, 96(2):175–194, jan. 2012.
 - [37] D. L. DONOHO et X. HUO : Uncertainty principles and ideal atomic decomposition. *Information Theory, IEEE Transactions on*, 47(7):2845–2862, nov. 2001.
 - [38] M. ELAD et M. AHARON : Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries. *Image Processing, IEEE Transactions on*, 15(12):3736–3745, déc. 2006.
 - [39] M. ELAD et A. FEUER : Super-resolution reconstruction of image sequences. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(9):817–834, sept. 1999.
 - [40] M. ELAD et Y. HEL-OR : A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. *Image Processing, IEEE Transactions on*, 10(8):1187–1193, août 2001.
 - [41] G. FACCIOLO, A. ALMANSA, J.-F. AUJOL et V. CASELLES : Irregular to Regular Sampling, Denoising, and Deconvolution. *Multiscale Modeling & Simulation*, 7(4): 1574+, 2009.
 - [42] S. FARSIU, M. ELAD et P. MILANFAR : Multiframe demosaicing and super-resolution of color images. *Image Processing, IEEE Transactions on*, 15(1):141–159, jan. 2006.
 - [43] S. FARSIU, D. ROBINSON, M. ELAD et P. MILANFAR : Advances and challenges in super-resolution. *Int. J. Imaging Syst. Technol.*, 14(2):47–57, 2004.
 - [44] S. FARSIU, D. ROBINSON, M. ELAD et P. MILANFAR : Robust shift and add approach to superresolution. In *SPIE Annual Meeting 2003*, vol. 5203, p. 121+. The International Society for Optical Engineering., jan. 2004.

- [45] S. FARSIU, M. D. ROBINSON, M. ELAD et P. MILANFAR : Fast and robust multi-frame super resolution. *Image Processing, IEEE Transactions on*, 13(10):1327–1344, oct. 2004.
- [46] R. FRANSENS, C. STRECHA et L. VAN GOOL : Optical flow based super-resolution : A probabilistic approach. *Computer Vision and Image Understanding*, 106(1):106–115, avr. 2007.
- [47] D. GEIGER, T.-L. LIU et M. DONAHUE : Sparse Representations for Image Decompositions. *International Journal of Computer Vision*, 33(2):139–156, 1999.
- [48] G. GOLUB et V. PEREYRA : Separable nonlinear least squares : the variable projection method and its applications. *Inverse Problems*, 19(2):R1+, avr. 2003.
- [49] G. GOLUB et C. VAN LOAN : *Matrix computations*, vol. 3. Johns Hopkins Univ Pr, 1996.
- [50] M. GRANADOS, B. AJDIN, M. WAND, C. THEOBALT, H. P. SEIDEL et H. P. A. LENSCH : Optimal HDR reconstruction with linear digital cameras. *In CVPR*, p. 215–222, 2010.
- [51] K. GRÖCHENIG et T. STROHMER : Numerical and theoretical aspects of nonuniform sampling of band-limited images. *In Nonuniform sampling : Theory and Practice*, Inf. Technol. Transm. Process. Storage, p. 283–324. Kluwer/Plenum, New York, 2001.
- [52] B. GUNTURK et M. GEVREKCI : High-resolution image reconstruction from multiple differently exposed images. *Signal Processing Letters, IEEE*, 13(4):197 – 200, april 2006.
- [53] R. C. HARDIE, K. J. BARNARD et E. E. ARMSTRONG : Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *Image Processing, IEEE Transactions on*, 6(12):1621–1633, déc. 1997.
- [54] F. J. HARRIS : On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE*, 66(1):51–83, jan. 1978.
- [55] R. HARTLEY et A. ZISSERMAN : *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, 2004.
- [56] S. W. HASINOFF, F. DURAND et W. T. FREEMAN : Noise-optimal capture for high dynamic range photography. *In CVPR*, p. 553–560, 2010.
- [57] Y. HE, K.-H. YAP, L. CHEN et L.-P. CHAU : A Nonlinear Least Square Technique for Simultaneous Image Registration and Super-Resolution. *Image Processing, IEEE Transactions on*, 16(11):2830–2841, nov. 2007.
- [58] M. IRANI et S. PELEG : Improving resolution by image registration. *CVGIP : Graphical Models and Image Processing*, 53(3):231 – 239, 1991.
- [59] A. A. KARATSUBA : An estimate of the L 1-norm of an exponential sum. 64(3):401–404, 1996.
- [60] K. KIRK et H. J. ANDERSEN : Noise characterization of weighting schemes for combination of multiple exposures. *In BMVC*, p. 1129–1138, 2006.

-
- [61] P. KUPPINGER, G. DURISI et H. BOLCSKEI : Uncertainty Relations and Sparse Signal Recovery for Pairs of General Signal Sets. *Information Theory, IEEE Transactions on*, 58(1):263–277, jan. 2012.
 - [62] M. LEBRUN, A. BUADES et J. MOREL : A nonlocal bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, 6(3):1665–1688, 2013.
 - [63] D. G. LUENBERGER et Y. YE : *Linear and Nonlinear Programming*.
 - [64] J. MAIRAL, F. BACH, J. PONCE, G. SAPIRO et A. ZISSERMAN : Non-local sparse models for image restoration. In *Computer Vision, 2009 IEEE 12th International Conference on*, p. 2272–2279. IEEE, 2009.
 - [65] F. MALGOUYRES et F. GUICHARD : Edge Direction Preserving Image Zooming : A Mathematical and Numerical Analysis. *SIAM Journal on Numerical Analysis*, 39(1), 2002.
 - [66] S. MALLAT et G. YU : Super-Resolution With Sparse Mixing Estimators. *Image Processing, IEEE Transactions on*, 19(11):2889–2900, nov. 2010.
 - [67] S. MANN et R. W. PICARD : On being ‘undigital’ with digital cameras : Extending dynamic range by combining differently exposed pictures. In *Proceedings of IS&T*, p. 442–448, 1995.
 - [68] A. MARQUINA et S. OSHER : Image Super-Resolution by TV-Regularization and Bregman Iteration. *J. Sci. Comput.*, 37(3):367–382, déc. 2008.
 - [69] A. MELMAN : Generalizations of Gershgorin disks and polynomial zeros. *Proceedings of the American Mathematical Society*, 138(7):2349–2364, 2010.
 - [70] P. MILANFAR : *Super-resolution imaging*, vol. 1. CRC Press, 2010.
 - [71] K. MITRA, A. VEERARAGHAVAN et R. CHELLAPPA : Analysis of Sparse Regularization Based Robust Regression Approaches. *Signal Processing, IEEE Transactions on*, 61(5):1249–1257, mars 2013.
 - [72] T. MITSUNAGA et S. K. NAYAR : Radiometric self calibration. In *CVPR*, p. 1374–1380, 1999.
 - [73] L. MOISAN : Periodic Plus Smooth Image Decomposition. *Journal of Mathematical Imaging and Vision*, 39(2):161–179, fév. 2011.
 - [74] H. NAKAI, S. YAMAMOTO, Y. UEDA et Y. SHIGEYAMA : High resolution and high dynamic range image reconstruction from differently exposed images. In *Proceedings of the 4th International Symposium on Advances in Visual Computing, Part II, ISVC ’08*, p. 713–722, Berlin, Heidelberg, 2008. Springer-Verlag.
 - [75] S. G. NARASIMHAN et S. K. NAYAR : Enhancing resolution along multiple imaging dimensions using assorted pixels. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(4): 518–530, avr. 2005.
 - [76] M. K. NG et N. K. BOSE : Mathematical analysis of super-resolution methodology. *Signal Processing Magazine, IEEE*, 20(3):62–74, mai 2003.

- [77] M. NIKOLOVA : A Variational Approach to Remove Outliers and Impulse Noise. 20(1-2):99–120, 2004.
- [78] Q. NING, K. CHEN, L. YI, C. FAN, Y. LU et J. WEN : Image Super-Resolution Via Analysis Sparse Prior. *Signal Processing Letters, IEEE*, 20(4):399–402, avr. 2013.
- [79] A. PAPOULIS : Generalized sampling expansion. *Circuits and Systems, IEEE Transactions on*, 24(11):652–654, nov. 1977.
- [80] S. C. PARK, M. K. PARK et M. G. KANG : Super-resolution image reconstruction : a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, mai 2003.
- [81] T. Q. PHAM, L. J. van VLIET et K. SCHUTTE : Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP J. Appl. Signal Process.*, 2006:236, jan. 2006.
- [82] J. S. PICARD et A. J. WEISS : Bounds on the Number of Identifiable Outliers in Source Localization by Linear Programming. *Signal Processing, IEEE Transactions on*, 58(5):2884–2895, mai 2010.
- [83] G. POPE, A. BRACHER et C. STUDER : Probabilistic Recovery Guarantees for Sparsely Corrupted Signals. *Information Theory, IEEE Transactions on*, 59(5):3104–3116, mai 2013.
- [84] M. PROTTER, M. ELAD, H. TAKEDA et P. MILANFAR : Generalizing the nonlocal-means to super-resolution reconstruction. *Image Processing, IEEE Transactions on*, 18(1):36–51, jan. 2009.
- [85] D. RAJAN et S. CHAUDHURI : An MRF-Based Approach to Generation of Super-Resolution Images from Blurred Observations. 16(1):5–15, 2002.
- [86] B. S. REDDY et B. N. CHATTERJI : An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8):1266–1271, août 1996.
- [87] E. REINHARD, G. WARD, S. N. PATTANAIK et P. E. DEBEVEC : *High Dynamic Range Imaging - Acquisition, Display, and Image-Based Lighting*. Morgan Kaufmann, 2005.
- [88] D. ROBINSON, S. FARSIU et P. MILANFAR : Optimal Registration Of Aliased Images Using Variable Projection With Applications To Super-Resolution. *The Computer Journal*, 52(1):31–42, jan. 2009.
- [89] D. ROBINSON et P. MILANFAR : Fundamental performance limits in image registration. *Image Processing, IEEE Transactions on*, 13(9):1185–1199, sept. 2004.
- [90] D. ROBINSON et P. MILANFAR : Statistical performance analysis of super-resolution. *Image Processing, IEEE Transactions on*, 15(6):1413–1428, juin 2006.
- [91] M. D. ROBINSON, C. A. TOTH, J. Y. LO et S. FARSIU : Efficient Fourier-Wavelet Super-Resolution. *Image Processing, IEEE Transactions on*, 19(10):2669–2681, oct. 2010.

-
- [92] A. SÁNCHEZ-BEATO : Coordinate-descent super-resolution and registration for parametric global motion models. *Journal of Visual Communication and Image Representation*, 23(7):1060–1067, oct. 2012.
 - [93] L. SBAIZ, P. VANDEWALLE et M. VETTERLI : Groebner basis methods for multichannel sampling with unknown offsets. *Applied and Computational Harmonic Analysis*, 25(3):277–294, nov. 2008.
 - [94] R. R. SCHULTZ et R. L. STEVENSON : Extraction of high-resolution frames from video sequences. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 5(6):996–1011, juin 1996.
 - [95] C. STUDER et R. G. BARANIUK : Stable restoration and separation of approximately sparse signals. *Applied and Computational Harmonic Analysis*, sept. 2013.
 - [96] C. STUDER, P. KUPPINGER, G. POPE et H. BOLCSKEI : Recovery of Sparsely Corrupted Signals. *Information Theory, IEEE Transactions on*, 58(5):3115–3130, mai 2012.
 - [97] H. SU, Y. WU et J. ZHOU : Super-Resolution Without Dense Flow. *IEEE Transactions on Image Processing*, 21(4):1782–1795, avr. 2012.
 - [98] H. TAKEDA, P. MILANFAR, M. PROTTER et M. ELAD : Super-Resolution Without Explicit Subpixel Motion Estimation. *Image Processing, IEEE Transactions on*, 18(9):1958–1975, 2009.
 - [99] G. TARTAVEL, Y. GOUSSEAU et G. PEYRÉ : Constrained Sparse Texture Synthesis. In A. KUIJPER, K. BREDIES, T. POCK et H. BISCHOF, édés : *Scale Space and Variational Methods in Computer Vision*, vol. 7893 de *Lecture Notes in Computer Science*, p. 186–197. Springer Berlin Heidelberg, 2013.
 - [100] J. TIAN et K.-K. MA : Stochastic super-resolution image reconstruction. *Journal of Visual Communication and Image Representation*, 21(3):232–244, avr. 2010.
 - [101] J. TIAN et K.-K. MA : A survey on super-resolution imaging. *Signal, Image and Video Processing*, 5(3):329–342, sept. 2011.
 - [102] Y. TRAONMILIN, S. LADJAL et A. ALMANSA : Robust Multi-image Processing With Optimal Sparse Regularization. *Preprint HAL url : <http://hal.archives-ouvertes.fr/hal-00940192>*.
 - [103] Y. TRAONMILIN, S. LADJAL et A. ALMANSA : Quantification de la robustesse de la super-résolution par minimisation L1. 2013.
 - [104] Y. TRAONMILIN et C. AGUERREBERE : Simultaneous High Dynamic Range and Super-Resolution Imaging Without Regularization. nov. 2013.
 - [105] Y. TRAONMILIN, S. LADJAL et A. ALMANSA : On the amount of regularization for super-resolution reconstruction. *Preprint HAL url : <http://hal.archives-ouvertes.fr/hal-00763984>*.
 - [106] Y. TRAONMILIN, S. LADJAL et A. ALMANSA : On the amount of regularization for super-resolution interpolation. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, p. 380–384. IEEE, août 2012.

- [107] Y. TRAONMILIN, S. LADJAL et A. ALMANSA : Outlier Removal Power of the L1-Norm Super-Resolution. *In Scale Space and Variational Methods in Computer Vision*, vol. 7893 de *Lecture Notes in Computer Science*, p. 198–209. Springer Berlin Heidelberg, 2013.
- [108] S. VAITER, C.-A. DELEDALLE, G. PEYRÉ, C. DOSSAL et J. FADILI : Local behavior of sparse analysis regularization : Applications to risk estimation. *Applied and Computational Harmonic Analysis*, déc. 2012.
- [109] P. VANDEWALLE, L. SBAIZ, J. VANDEWALLE et M. VETTERLI : Super-resolution from unregistered and totally aliased signals using subspace methods. *Signal Processing, IEEE Transactions on*, 55(7):3687–3703, juil. 2007.
- [110] P. VANDEWALLE, S. SÜ, M. VETTERLI *et al.* : A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP Journal on Advances in Signal Processing*, 2006, 2006.
- [111] E. VURAL et P. FROSSARD : Analysis of Descent-Based Image Registration. *SIAM Journal on Imaging Sciences*, 6(4):2310–2349, jan. 2013.
- [112] N. A. WOODS, N. P. GALATSANOS et A. K. KATSAGGELOS : Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images. *Image Processing, IEEE Transactions on*, 15(1):201–213, jan. 2006.
- [113] K.-H. YAP, Y. HE, Y. TIAN et L.-P. CHAU : A Nonlinear ℓ_1 -Norm Approach for Joint Image Registration and Super-Resolution. *Signal Processing Letters, IEEE*, 16(11):981–984, nov. 2009.
- [114] G. YU, G. SAPIRO et S. MALLAT : Solving Inverse Problems With Piecewise Linear Estimators : From Gaussian Mixture Models to Structured Sparsity. *Image Processing, IEEE Transactions on*, 21(5):2481–2499, mai 2012.
- [115] Q. YUAN, L. ZHANG et H. SHEN : Multiframe Super-Resolution Employing a Spatially Weighted Total Variation Model. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(3):379–392, mars 2012.
- [116] C.-H. ZHANG et T. ZHANG : A General Framework of Dual Certificate Analysis for Structured Sparse Recovery Problems, avr. 2012.
- [117] W. ZHAO et H. SAWHNEY : Is Super-Resolution with Optical Flow Feasible ? *In A. HEYDEN, G. SPARR, M. NIELSEN et P. JOHANSEN, édés : Computer Vision ECCV 2002*, vol. 2350 de *Lecture Notes in Computer Science*, chap. 40, p. 599–613. Springer Berlin Heidelberg, Berlin, Heidelberg, avr. 2002.
- [118] H. ZIMMER, A. BRUHN et J. WEICKERT : Freehand HDR imaging of moving scenes with simultaneous resolution enhancement. *Comput. Graph. Forum*, 30(2):405–414, 2011.

Relations entre le modèle d'image et le nombre de mesures pour une super-résolution fidèle

Yann Traonmilin

Résumé :

La super-résolution multi-image consiste à produire une image de haute résolution d'une scène à partir de plusieurs acquisitions de basse résolution. Le problème inverse correspondant peut être mal posé, il faut alors avoir recours à un modèle de régularité sur le contenu de la scène pour produire une image de haute résolution réaliste. Seulement, les erreurs de modélisation limitent la performance de telles méthodes. En conséquence, il est important de déterminer les conditions dans lesquelles le problème est bien posé afin d'éviter de régulariser lorsque cela est possible, et ainsi maximiser la fidélité de la super-résolution. On se pose donc les questions suivantes :

Pour des bruits d'énergie finie ou une contamination par des données aberrantes (de support fini), combien d'images permettent de produire une image de haute résolution fidèle à la réalité ? Comment maximiser la fidélité du résultat des méthodes régularisées lorsque le nombre d'images est insuffisant ?

Dans le cas d'un bruit de mesure, une étude du comportement asymptotique du conditionnement du système inverse valide la possibilité d'effectuer une super-résolution non régularisée lorsque le nombre d'images est suffisant. Par ailleurs, dans les cas proches du seuil critique d'inversibilité, qui sont les plus mal posés, nous proposons et validons un estimateur local du conditionnement qui nous permet de restreindre le plus possible le recours à une régularisation. Dans le cas de données aberrantes, nous utilisons l'équivalence entre le problème de reconstruction parcimonieuse et la résistance aux données aberrantes pour démontrer des bornes de résistance aux données aberrantes du problème de super-résolution. Nous traitons aussi le cas de la régularisation et montrons les conditions sous lesquelles celle-ci permet d'améliorer la robustesse du problème. Tous ces résultats sont validés par des expériences.

Mots-clefs : super-résolution, régularisation, interpolation, données aberrantes, robustesse, modèles parcimonieux

Abstract : Multi-image super-resolution produces a high resolution image from several low resolution acquisitions of a scene. This inverse problem can be ill-posed. We then need to use a regularity model on the scene to be able to produce a realistic image. However, modelization errors limit the performance of such methods. Consequently, finding conditions where the problem is well-posed is necessary to be able to limit the amount of regularization when possible, and maximize the fidelity of the result. We ask ourselves the following questions :

For noises with finite energy or outliers, how many images permit the reconstruction of a high resolution image close to the real scene ? How to maximize fidelity of regularized methods when the number of images is too small ?

For measure noises, an asymptotic study of the conditioning guarantees that it is possible to use an unregularized method if enough images are available. In cases closes to the critical inversible case, which are not well-posed, we propose and validate a local estimator of the conditioning, which we use to limit the amount of regularization. For outliers, we use the equivalence between the sparse recovery problem and the robustness to outliers to calculate bounds for the robustness of super-resolution. We also study the regularized case and show conditions which increase the robustness of the problem. All these results are validated by experiments.

Keywords : super-resolution, regularization interpolation, outliers, robustness, sparse models